Developing Human Stem Cell Derived Motor Neuron Models of Amyotrophic Lateral Sclerosis

Developing Human Stem Cell Derived Motor Neuron Models of Amyotrophic Lateral Sclerosis

A dissertation presented

by

Jackson L. Sandoe

to

The Department of Molecular and Cellular Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biochemistry

Harvard University

Cambridge, Massachusetts

August, 2014

Dissertation Advisor: Kevin Eggan Jackson L Sandoe

Developing Human Stem Cell Derived Motor Neuron Models of Amyotrophic Lateral Sclerosis

Abstract

Human neurodegenerative disorders are among the most difficult to study. In particular, the inability to readily obtain the faulty cell types most relevant to these diseases has impeded progress for decades. Amyotrophic lateral sclerosis is a late onset neurodegenerative disease in which the upper and lower motor neurons degenerate, leading to paralysis and eventually death. Recent advances in pluripotent stem cell technology now grant access to significant quantities of disease pertinent neurons both with and without predisposing mutations. The two studies described in this thesis demonstrate the feasibility of using MNs, generated from pluripotent stem cell lines harboring known ALS mutations, to establish in-vitro models of the disease. Specifically, we first used gene targeting to establish genetically controlled systems, able to identify causal relationships between a familial ALS mutation and in vitro phenotypes. Next, using transcriptional profiling, we identified novel pathways altered by the mutation and demonstrated functional consequences of these pathways' misregulation. Furthermore, by monitoring the physiology of the pluripotent stem cell derived MNs, we discovered an increased firing rate in the mutant MNs, and identified an FDA-approved drug, retigabine, capable of rescuing this defect. Lastly, to aid in the discovery of additional therapeutic compounds, we combined gene targeting, transcriptional profiling, and a fluorescent reporter human embryonic stem cell line to establish a well-controlled in vitro system capable of identifying genetic modifiers of the phenotypes described herein.

Table of Contents

Acknowledgements

First I would like to thank my mentor, Kevin Eggan, for allowing me to pursue my research interests and providing thoughtful advice throughout my Ph.D. Also, I would like to thank all past and present members of the Eggan lab for providing a working environment which was both enjoyable and intellectually stimulating. I could not have performed the many hours of research required to complete my Ph.D. without friends to push me along. I have also valued the advice of my dissertation committee over the years. I would also like to thank my future wife, Sunny, who has supported me and aided in numerous aspects of my career.

Chapter 1: Introduction

The majority of this Chapter was previously published as

Sandoe J, Eggan K. Opportunities and challenges of pluripotent stem cell neurodegenerative disease models. *Nature Neuroscience*. 2013 July;16(7):780-9. PMID: 23799470

ALS as a Model Pathology for Disease Modeling Experiments

For historical as well as methodological reasons, the application of patient-specific stem cells in disease modeling experiments is a particularly promising approach for ALS. First, a vast amount of data has been generated to describe the disorder. The clinical presentation of ALS has been characterized for over 100 years, as have the loss of spinal and cortical MNs, and the typical progression of the disease into paralysis and ultimately death[1]. The disease can present as bulbar or limb onset, but commonly progresses to impacting the entire motor system[1]. Also thoroughly described are the results of post-mortem sections, which often reveal ubiquitin and other protienaceous inclusions in the spinal cord[1]. This clinical data establishes common pathological findings which can be used as starting points for in vitro disease modeling experiments.

Second, stem cell disease modeling experiments are a promising means for parsing the complex set of genetic and environmental factors that may contribute to the development of the disorder. While the predominant patient group develops the disease in a seemingly sporadic manner (sALS), 17%-23% of patients with ALS report a family history of the disorder and are classified as familial ALS (fALS) patients[2-4]. Dominant inheritance is the most commonly reported transmission pattern, although recessive inheritance has been documented in some cases[1]. Family predisposition to the disorder has led to many genetic studies of affected genealogies, leading to the discovery of over 20 familial ALS mutations[1]. The most common mutations are found in the C9orf72, SOD1, FUS, and TDP-43 genes[1]. While most mutations are missense amino acid changes, the C9orf72 gene was found to contain a large hexanucleotide expansion[5]. The diverse function of the genes which when mutated can lead to ALS has not enabled a clear pathway of disease to be established. Furthermore, every gene which has been implicated in fALS has also been found to be mutated in a smaller fraction of

sALS cases[1]. The implication of fALS genes in sALS is further supported by pathological findings of TDP-43 and to a lesser degree SOD1 positive inclusion in post mortem section of sALS spinal motor neurons[6,7]. While some of these sALS cases may have de novo mutations within the genes more commonly associated with fALS, it is also likely small family size, misdiagnosis of ALS, and incomplete penetrance has resulted in an underrepresentation of fALS. Further complicating the study of ALS in patients is the impact of environmental factors or other genetic modifying mutations which can influence the clinical representation of ALS[1]. Supporting this conclusion is the large patient heterogeneity which is seen even between individuals which harbor the same fALS mutation[1,3]. Patient groups with the same fALS mutations can present with variable disease onset, progression of the disorder, and bulbar or limb onset[3]. A more specific example of patient heterogeneity is found in patients identified as carriers of the C9orf72 hexanucleotide expansion, who have been shown to present symptoms of ALS, Frontal Temporal Dementia (FTD), or a combination of both ALS and FTD[8]. These findings have dispelled the notion that ALS may be a 'many genes, one degenerative syndrome' and instead implicates ALS on a spectrum of neurodegenerative disorders which is impacted by patient genetics and environmental factors[1]. The complicated nature of patient biology has led many to the use of model systems for the study of ALS, in the hopes that reducing the unknown factors which can modulate the disease may facilitate the discovery of disease mechanisms.

Finally, recent results from studies employing animal models of ALS have pointed toward the need for human models. Multiple animal models of ALS have been established, and mouse models have been utilized particularly extensively[9]. Mice which express the human SOD1 gene harboring known familial mutations present many signs of ALS commonly found in human patients[9]. Therefore, a considerable amount of work has gone into uncovering a multitude of phenotypes associated with the mutations at various points during disease progression. Some major findings have been the induction of caspase activity, formation of protein inclusions, a major contribution of glial cells to MN toxicity, and the early breakdown of

the motor unit leading to de-innervation of the muscle[9-11]. While the mouse models have been informative, they rely on extreme overexpression of the mutant SOD1 transgene, as lower expressing mouse models either have large phenotypic variability or do not succumb to any measurable defect, and thus any conclusions maintain the caveat of pathological artifacts possibly being driven by protein overexpression[12]. Unfortunately, many drugs which have proven effective in the mouse model have failed in human trials, pointing toward the need for a human model[13]. To address the desire for human models of ALS pathology, I have worked to establish in vitro systems using MNs derived from human pluripotent stem cells. In this chapter I will discuss the impediments I have encountered in using stem cells to understand neurodegeneration and offer strategies to surmount them. In particular, I'll focus on areas that I feel have not yet been widely discussed. For instance, while significant effort has been put forth to describe the potential distinctions between hiPSCs and hESCs, there has been little discussion concerning the ramifications that differences between any two pluripotent stem cell lines might have on the outcome of a disease modeling experiment. Additionally, although substantial resources have been invested in methods for directed differentiation of stem cells into specific neural types and their functional characterization at early time points, little is known concerning how these neurons change, or mature over time in culture. These and other sources of variation between pluripotent cell lines and within neuronal cultures may result in misleading red herring findings or could obscure real disease related discoveries and are veritable challenges that must be surmounted.

The Opportunities of Pluripotent Stem Cells

The opportunity to produce discrete populations of neuronal types from human pluripotent cell lines has inspired a new approach to the study of human neurodegeneration. With workable methods for stem cell differentiation in hand, it is possible to generate human neurons in substantial numbers for monitoring and studying disease processes[14]. The last piece

of enabling technology came in the landmark discovery of reprogramming methods using defined factors[15] and the subsequent derivation of the first patient specific human induced pluripotent stem cell (hiPSC) lines[16,17]. Now, by using hiPSCs, it has become possible to produce neurons carrying the precise constellation of genetic variants that caused neurodegeneration in a given individual[16]. These neurons enable the examination of processes not easily observed in postmortem tissues, including important events that might occur prior to disease onset[18]. Continued refinement and deployment of this "neurodegeneration-in-a-dish" approach holds substantial promise for gathering new insights into diseases of the nervous system. Indeed, many disease related phenotypes are beginning to be documented from patient derived hiPSCs[14,19]. For example, changes in the processing of proteins implicated in ALS, Parkinson's, Huntington's, and Alzheimer's diseases have been found in neurons produced from patient derived iPSCs[20-22]. Additionally, findings reminiscent of terminal stage phenotypes, such as cell death, have been observed in early onset disorders including Spinal Muscular Atrophy and Lesch-Nyhans[23,24]. Furthermore, phenotypes implicated in behavioral disorders such as neural connectivity and synapse formation have been observed using hiPSCs derived from patients diagnosed with schizophrenia and Rett syndrome[25,26]. While these early studies are important demonstrations in the utilization of hiPSC technology, their success may be the exception rather than the rule given our current understanding of the variability underlying many of these experiments. We now know of a number of obstacles that are routinely encountered by investigators employing these technologies, which we have come to realize must be carefully considered and properly overcome if accurate conclusions are to be drawn[27]. Attention to these would-be pitfalls will help unlock the full potential of this approach to understanding neurological disease.

From the Embryo to Patient Reprogramming

Inspired by progress in the study of mouse stem cells, early work on human pluripotency

enabled the establishment of favorable conditions for both the routine derivation of hESC lines from discarded IVF embryos and the support of their long-term self-renewal[28-30]. The hESC lines that resulted from those efforts, mostly derived from embryos presumed to have a normal health status, allowed preparations of human neurons to be produced and in turn used to interrogate environmental or non-cell autonomous factors implicated in disease[31,32]. While these "normal" hESC lines have been incredibly useful, it seemed a logical progression of this approach to produce hESCs that harbored disease-predisposing mutations and study neuronal populations generated from them[33]. Although disease-specific hESC lines were subsequently derived from embryos after pre-implantation genetic diagnosis (PGD), the scarcity of the donated material needed, and the limited number of conditions for which PGD was routinely performed, prevented the main-stream implementation of this approach[34,35]. Today, methods for transcription factor mediated reprogramming have largely circumvented the need for PGD embryos by allowing almost any patient tissue available to be used for the production of hiPSC lines[15,35]. Such readily adoptable methods for deriving stem cell lines have now been combined with existing approaches for neuronal differentiation in a number of early attempts to model neurological disease, which have been comprehensively reviewed previously[19,36]. While the advent of reprogramming proved to be a critical enabler, it raised significant questions concerning whether hiPSCs were significantly different from hESCs[37]. Although the resolution of this question is important, it has perhaps in retrospect distracted from a larger issue with more significant ramifications: How different are the properties of any two hiPSC lines produced from the same person, or for that matter, from multiple individuals? Furthermore, are the effects of these idiosyncratic differences between lines greater than the effects I can reasonably hope to observe as a result of any disease causing DNA variant I plan to study? Although this variability between pluripotent cell lines is still poorly understood, it has now begun to be recognized as a potential confounding factor in studies[27]. As a result, movement has begun towards the replacement of qualitative assays for pluripotency, such as immunostaining, or teratoma

formation with more quantitative measures of a given cell-line's performance[38,39].

Sources of Variability

Each pluripotent stem cell-line, whether a hESC or hiPSC, has its own particular history. Whether a given cell line originated from an embryo or somatic cell, the method of derivation that was employed and the subsequent culture medias or environments to which it has been exposed, have all been proposed to contribute to its inherent properties[29,40,41]. As a result, it is conceivable that these sources of variability could interfere with the comparison of any two cell lines in the context of disease related experiments. Consequently, careful consideration should be given to the cell lines that are selected for any given study and how their quality is controlled. Historically, the pluripotency of a cell-line was verified by qualitative methods such immunostaining with antibodies recognizing antigens selectively expressed in stem cells, or by injection into nude mice to assess their capacity to form benign teratomas [30,42]. While these assays were useful for verifying whether a cell line was indeed pluripotent, they fell short of providing information concerning which lines were likely to be useful for a particular application[42]. More recently, genome wide analyses of transcription and DNA methylation in larger collections of pluripotent cells have begun to catalog potential sources of variation and attempt correlations to a given line's differentiation propensity or functional behaviors[27,39]. These studies have been instrumental in both motivating the reevaluation of previous metrics for pluripotency and the implementation of new measures.

Many groups have now compared hiPSCs with a few well-established hESC lines and have suggested that differences between the two classes of cell lines in gene expression and methylation are present[41,43]. Some studies have concluded the transcriptional distinctions between stem cell-types originate from epigenetic memory of the somatic cell from which the iPS cells were derived[41,44]. However, others have suggested that this memory is rapidly lost after modest passage of established lines[45]. When our lab examined 20 hESC lines and a

dozen hiPSC lines, in one of the largest studies to date, we too found substantial variability in transcription and DNA methylation between the many hESC and iPS cell lines[39]. While we found that on average, hESCs and hiPSCs were modestly distinguishable, there was a considerable amount of overlap between the two groups[39] (Figure 1.1A). Furthermore, the average difference between any two human hESCs or any two hiPSCs was often greater than the average differences between the two classes[39]. As a result, we concluded that there are many iPSCs that are more like the average hESC then any particular hESC might be[39]. In other words, although there has been a good deal of focus on the inherent variation between iPSC and hESCs, the differences that exist between *any* two lines selected, regardless of how they were made, are likely to be greater. Although it may or may not be that variability of transcription in the pluripotent state leads to similar variation in differentiated neurons, it is well known that changes in DNA-methylation are mitotically heritable[41]. Moreover, it has been shown that cell-line variation can impact the phenotype of differentiated cells and neurons produced from a given cell line[27,39].

Recently, the Eggan lab and other groups realized that a significant component of the transcriptional and DNA methylation variation between cell lines is absent from male pluripotent cells[24,40]. Further studies suggest this variation results from a gradual erosion of dosage compensation leading to an increase in X-linked gene expression, which was not corrected by neuronal differentiation[24]. As there are dozens of X-linked genes with nervous system functions, modest changes in the expression of these genes could be significantly confounding co-variables in many assays of neuronal function. In particular, we demonstrated perturbed dosage compensation had a substantial impact on the study of X-linked *HPRT* mutations, which cause Lesch-Nyhan syndrome[24]. In light of these findings, it may be wise to either use male cells lines, or monitor the extent to which female pluripotent cell lines have retained an inactive X. We find that staining cultures with antibodies identifying the inactive X, such as those recognizing H3-K273me proves to be an effective strategy[46]. The use of assays for monitoring

proper maintenance of X-inactivation will be particularly important when modeling X linked disorders. More generally, it will be increasingly important to understand how epigenetic variation arises, and where any particular cell line selected for disease related studies falls within the range of variability. The use of increasingly established profiling assays such as "Pluritest" may then allow some deviant lines to be identified and discarded[38].

On a more encouraging note, while differences in derivation method and culture history cause significant variation, they may not completely overcome the subtle effects of genetic background. Chen et al (2009) established a large set of hES cells including a considerable cohort of sibling pairs and analyzed them by transcriptional profiling[29]. Hierarchal clustering of the resulting data grouped sibling pluripotent stem cell lines together, regardless of how the cell lines were derived, suggesting that changes in genetic background may be detected transcriptionally despite environmental sources of variability[29]. However, this experiment raises another issue, the potential importance of genetic background when evaluating or selecting cell lines for disease modeling experiments.

While a thorough study of transcriptional and epigenetic variance within stem cell lines has provided some insight, the catalogs of genome-wide information that have resulted will be most informative when they are compared with the results of quantitative assays that probe the actual behavior of a stem cell line and its derivatives[39]. Early quantitative studies of stem cell differentiation support the view that each line displays individual propensities for differentiation down a given lineage[27,47] (Figure 1.1B). If not taken into account this discordant behavior could cause significant problems when evaluating neurological disease modeling experiments. For example, it has been shown that many stem cell lines are more difficult to neuralize and therefore produce fewer neural progenitors and post-mitotic neurons[27,48]. Therefore, if lines such as these were chosen within a cohort used for studies, it could lead to erroneous assignment of a phenotype, merely as a result of chance. One approach proposed for overcoming this dilemma is to select lines by simply and rapidly assessing their differentiation propensity using a

"lineage scorecard"[39]. By allowing stem cell lines to differentiate without the influence of small molecules or growth factors, they generate progenitor and terminally differentiated cell types in an unconstrained and random manner. By then measuring transcription of a few key genes with expression patterns demarcated to specific lineages, a reproducible and quantitative representation of cell line early differentiation propensities can be obtained[39]. Impressively, it was found that this approach was able to very accurately predict which lines were most useful for producing neurons for disease related studies[27,39]. It is likely that such "lineage scorecards" or other simple, quantitative measures of cell line behavior will become standard practice. These assays will allow investigators to rapidly discard cell lines that behave poorly and optimally select those with a similar capacity to produce neuronal-types of interest, harmonizing downstream assays. The only risk I can foresee from such assays is that they could in principal lead to the elimination of an unanticipated effect that a given genetic variant might have on early development. If alterations in neuronal progenitor activity or development are implicated in the disorder being studied, the elimination of pluripotent cell lines with poor neuronal differentiation propensity may mask disease specific progenitor phenotypes by preferentially excluding the patient derived pluripotent cell lines which would display the most severe developmental defects. However, if mature neurons are to be investigated because the disorder under study is late onset, the benefits realized from such assays for the production of reproducible populations of differentiated neuronal types that could be then rigorously assessed for disease related phenotypes seems to far outweigh these risks.

Figure 1.1

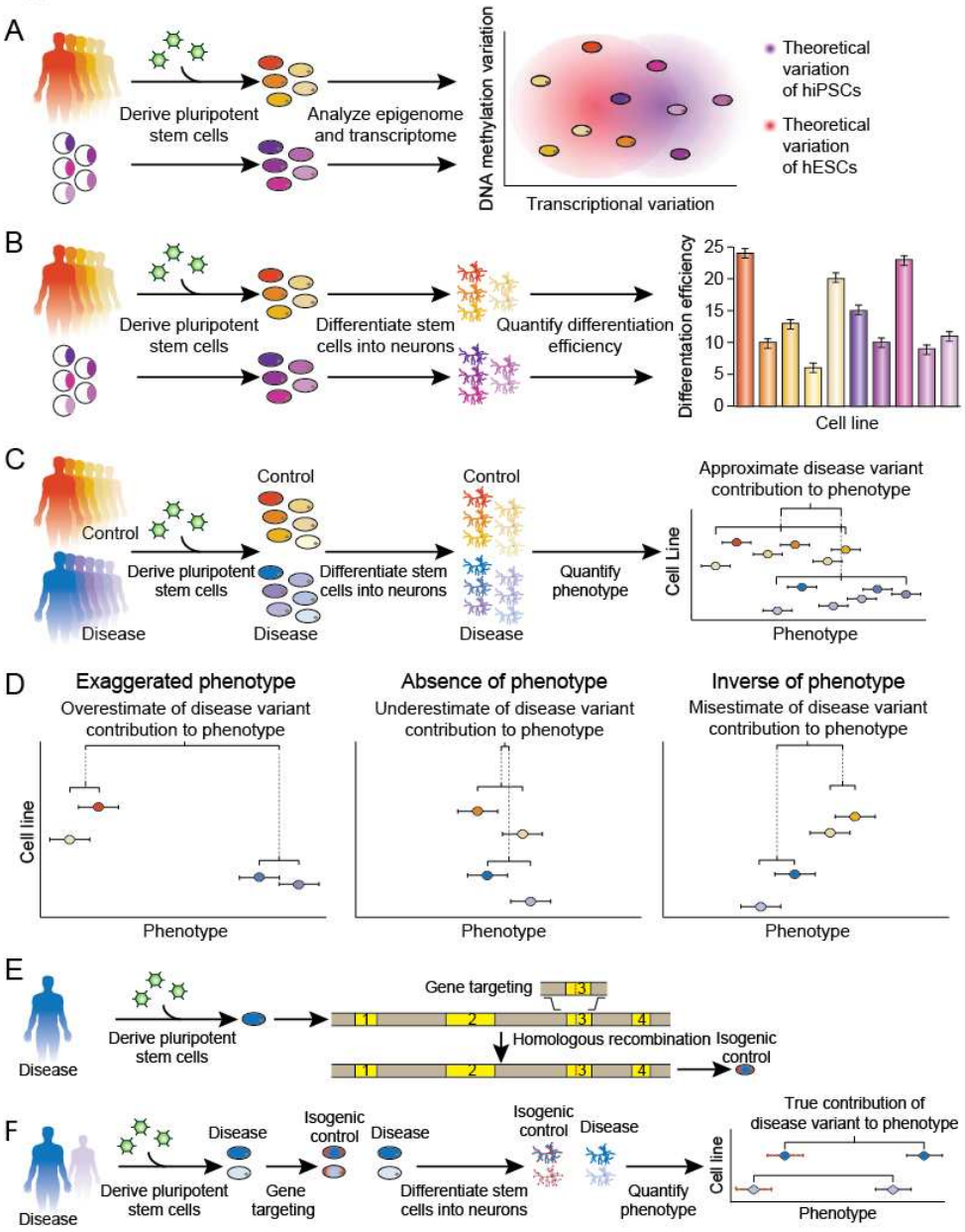
A
Derive pluripotent stem cells
Analyze epigenome and transcriptome
DNA methylation variation
Transcriptional variation
Theoretical variation of hiPSCs
Theoretical variation of hESCs
B
Derive pluripotent stem cells
Differentiate stem cells into neurons
Quantify differentiation efficiency
Differentation efficiency
Cell line
C
Control
Disease
Derive pluripotent stem cells
Differentiate stem cells into neurons
Quantify phenotype
Approximate disease variant contribution to phenotype
Cell Line
Phenotype
D
Exaggerated phenotype
Overestimate of disease variant contribution to phenotype
Absence of phenotype
Underestimate of disease variant contribution to phenotype
Inverse of phenotype
Misestimate of disease variant contribution to phenotype
Cell line
Phenotype
E
Disease
Derive pluripotent stem cells
Gene targeting
Homologous recombination
Isogenic control
F
Disease
Derive pluripotent stem cells
Gene targeting
Isogenic control
Differentiate stem cells into neurons
Quantify phenotype
True contribution of disease variant to phenotype
Cell line
Phenotype

Figure 1.1 (Continued) Variability amongst cohorts of pluripotent stem cells may lead to erroneous conclusions in disease modeling experiments driven by atypical cell lines. (A) Pluripotent stem cell lines derived from reprogrammed fibroblasts (hiPSC) or discarded embryos (hESC) demonstrate a range of unique methylation and gene expression profiles. (B) Studies have also demonstrated a considerable amount of variation in differentiation propensity amongst pluripotent cell lines. (C) Although few studies have systematically interrogated neurons derived from a large cohort of cell lines in phenotypic assays, it is likely that each individual cell line will produce neurons that fall within a broad range of the measured phenotype. (D) This variability necessitates the use of multiple control and patient-derived hiPSC lines in order to rule out the possibility that atypical cell lines are driving or masking statistically significant cell survival phenotypes. (E) Gene targeting techniques allow the generation of isogenic cell lines which differ at only the disease-associated variant being studied. (F) The use of neurons derived from isogenic pluripotent cell lines will reduce much of the variability between control and patient-derived neurons and highlight contributions of the variant to neuronal phenotypes. (All graphs are hypothetical and not from actual data.)

Overcoming the Noise

Several approaches have been utilized to account for the variability among human pluripotent stem cell lines in modeling experiments. First, some studies have used large cohorts of both patient and control pluripotent stem cell lines to derive neurons[49,50] (Figure 1.1C). By comparing neurons differentiated from many distinct patients and controls, plausible disease relevant phenotypes within a patient cohort have been reported[50]. This approach is perhaps most convincing when multiple groups carry out similar phenotypic assessments, providing a more rigorous platform for interrogating the same cell lines as has been reported[50]. While attempting to correlate any particular finding with cases or controls in a given study can be informative, it cannot demonstrate that the finding in question is the direct result of the known causative mutation, or mutations, the disease cases in question may harbor. Such studies leave open the formal possibility that the findings are accidental and the result of unrelated and poorly understood variation such as those described above (Figure 1.1D). Another weakness of this approach is the potentially large, and unknowable, number of cell lines that might be necessary to convincingly discover small differences between patients and controls. This is a particularly important consideration when investigating subtle phenotypes, which require more cell lines in order to ensure generalizability of results. In the case of extremely subtle phenotypes, examining the number of cell lines necessary to reach sufficient confidence in the reliability of

group differences may become impractical and cost-prohibitive.

One proposed solution to this problem is to utilize pluripotent stem cells from otherwise genetically related individuals differing in their disease related phenotype[23]. While this approach might successfully reduce genetic background noise, new problems could be introduced if this strategy were used for modeling idiopathic or sporadic forms of a disease. For instance it could be that undiagnosed but genetically vulnerable individuals participating in such studies could develop the disease after the time of iPS cell derivation, inappropriately skewing findings in a cohort of control cell lines. I conclude that while comparisons between neuronal populations derived from cases and controls may allow preliminary identification of disease related phenotypes, whenever possible, they should be married with genetic complementation studies designed to test the direct relationship between phenotype and genotype. In some cases, the use of viral or transgenic methods to overexpress or reduce the transcriptional levels of genes implicated in disease progression may provide substantial evidence for their role in observed pathology[51]. However, advances in gene targeting technology now offer more elegant approaches to genetic complementation studies.

Using gene targeting technology to remove mutations from pluripotent stem cell lines can create isogenic cell lines that differ at only one base pair[52,53]. This approach should in principal eliminate variation arising from stem cell line derivation and genetic background, thereby producing more ideal disease and control cell lines (Figure 1.1E and F). While many methods have been used in the past for gene targeting in pluripotent stem cell lines (Bacterial Artificial Chromosome[54], Adeno Associated Virus[55], Helper Dependent Adeno Virus[56], single stranded Oligonucleotides[57], nuclease driven gene targeting is rapidly becoming established as a dominant approach[53,58,59]. By generating site-specific nucleases, a double strand break (DSB) can be induced at a pre-designated locus in the genome[60]. This DSB is often repaired by nonhomologous end joining, an error-prone process, which can generate frame-shift mutations[60]. However, if a plasmid is simultaneously introduced that contains homology to the

sequence in which the DSB is induced, the homologous sequences can be used as a template for repair by homologous recombination[60]. Nuclease mediated homologous recombination has been used successfully to both correct and introduce known disease causing mutations[61,62]. Additionally, cDNAs can be targeted to loci resistant to transgene silencing, allowing for overexpression of cDNAs in an isogenic setting[63].

The primary concern surrounding nucleus-mediated gene-targeting is that these nucleases may induce collateral and unanticipated mutations at "off-target" sites around the genome. Exome and whole genome sequencing have been used to search for such off-target mutations and thus far they have been found only at a relatively low frequency[64,65]. However, mutations induced by nuclease activity are not the only possible risk inherent in this approach. In fact, the mutations derived from clonal selection are likely of as much concern as off-target nuclease activity[65]. Because of the risks associated with clonal isolation steps, it is advisable to generate multiple clones during gene targeting experiments, which can be used to verify any subsequent phenotypes found. While these strategies are becoming more advanced, common sense must still employed in their utilization. Proper examination of transcription and differentiation variability should still be performed before generation of isogenic pluripotent cell lines to assure the parental cell line is not atypical from the start and can efficiently produce the cell type of interest.

To Induce or Correct a Mutation, That is the Question

While gene targeting promises to allow the introduction of disease alleles into hES cells, this does not preclude the need for patient-specific iPS cells. In fact, many experiments critical for identifying and understanding relationships between genetic mutations and disease phenotypes are likely to continue to require patient-specific iPS cell lines[26]. For example, patient-derived iPS cells remain an interesting opportunity for modeling of neurological disorders with a strong genetic basis but for which specific disease-causing mutations are unknown[26].

Such a strategy could capture the constellation of variants that cause disease in an individual and allow their effects to be studied in disease affected neuronal populations[14]. Detection of a reproducible phenotype in such individuals could implicate the presence of genetic variants worthy of pursuit through genome sequencing. Also, the scalability of iPS cell derivation and the paucity of social concerns with this approach have resulted in the accumulation of a vast number of iPS cell lines[14,17]. This large number of pluripotent stem cell lines may one day facilitate the study of discrete disease causing variants in patients of multiple ethnicities, allowing the influence of genetic background to be explored and eventually the efficacy of potential therapeutics to be most broadly evaluated[14].

An additional and previously un-discussed benefit of combining gene-targeting with patient-derived iPS cells is the knowledge that protective genetic modifiers of a disease-causing variant in question are unlikely to confound *in vitro* studies. For example, as a multitude of novel disease variants are uncovered by genome-wide association studies, the use of pluripotent stem cell disease modeling approaches will be an attractive means to establish relationships between these mutations and neural dysfunction. It might initially seem attractive to introduce these mutations into a single "normal" genetic background, for instance into an hESC line, to study their influence on neuronal behavior. However, protective alleles on that genetic background might obscure phenotypes relevant to disease biology. Generation of isogenic controls by rescuing mutations found in patient-derived iPSCs would ensure that the important combination of modifying alleles for disease manifestation are present in the background being studied.

Differentiation of Pluripotent Cell Lines

After carefully establishing and evaluating pluripotent stem cell lines for use in disease modeling experiments, differentiation of these cell lines into the cell types of interest must next be pondered. Protocols for the production of many neural and nervous system cell-types are

either under development or have now been reported[66]. However, before embarking on disease related studies, it will continue to be important to thoroughly evaluate a given protocol's ability to generate sufficient quantities of the appropriately functional neuronal type. Immunocytochemistry and transcriptional profiling are often useful for rapidly assessing whether gene products found in particular neural types are present following stem cell differentiation[67]. However the lack of gene expression signatures that allow for the identification of the neuronal type of interest remains a problem for many experiments[68]. In these cases it may be necessary, and in perhaps all cases desirable, to use functional assays to determine whether the appropriate nervous system cell types have been produced. Examples of successful implementation of this approach include demonstration of neuromuscular junction formation by spinal motor neurons[67], capsaicin responsiveness in nociceptors[69], peptide secretion by neurosecretory cells[70] and ultimately successful engraftment after transplantation[71].

While confirming the presence and appropriate functionality of the central cell type of interest is essential, it is also important to characterize the other cells produced during the differentiation process. Most protocols result in the production of a myriad of cell types, only a small fraction of which will be the intended neural subtype[72] (Figure 1.2A). These heterogeneous cultures can cause complications in the interpretation of downstream experiments and should be taken into consideration when designing any disease modeling studies. Many neurological disorders have non-cell autonomous contributors to the disease which pose obstacles and limit conclusions that can be drawn from cultures containing diverse neural cell types[10,73] (Figure 1.2C). If a toxic cell type happens to also be produced during a given differentiation strategy, any phenotypic differences uncovered between disease and control neurons may need to be subsequently queried for its cell autonomy or non-cell autonomy. Modifications of protocols can alter the neuronal subtypes produced and it may be possible to find conditions which will yield the cell type of interest while limiting the production of other known toxic cell types[74]. Another

common component of cultures derived from human pluripotent stem cells are residual neural progenitors, which can present a substantial obstacle to detailed neuronal phenotyping, which I will discuss in detail below[72].

Figure 1.2

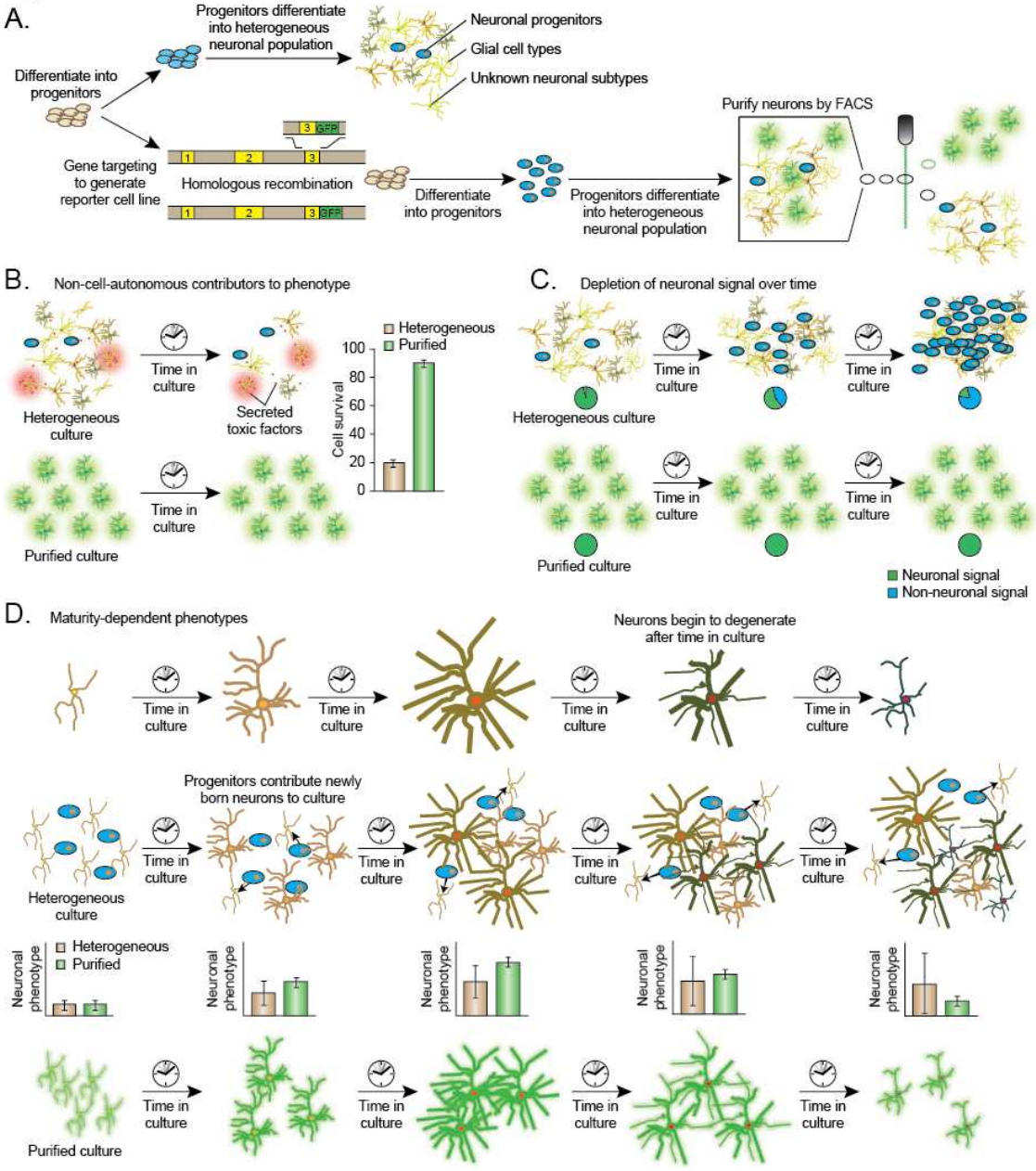

Figure 1.2 (Continued) The diversity of neuronal cultures derived from pluripotent stem cells can interfere with phenotypic observations. (A) Most differentiation protocols produce a heterogeneous population of neuronal progenitors, neurons and glial cell types. The use of reporter pluripotent stem cell lines enables the purification of neurons from heterogeneous populations by fluorescence-activated cell sorting (FACS). Heterogeneous cultures can impede experiments in many ways, most of which can be overcome by the purification and plating of a defined neuronal culture. (B) First, toxic cell types that contribute to non-cell-autonomous neuronal death may be produced during differentiation and may occlude mechanistic studies of cell survival phenotypes. (C) Second, the self-renewing neuronal progenitor cells can expand over time and greatly outnumber the postmitotic neurons, thus washing out possible neuron-specific disease signals. (D) Last, many neurodegenerative diseases are late onset and therefore may require a considerable time in culture for phenotypes to manifest. Thus, the introduction of newly born neurons into cultures by lingering progenitor cells can add considerable variability to phenotypes that depend on long durations in vitro. (All graphs are hypothetical and not from actual data.)

Needles in a Haystack

One approach to reducing phenotypic variation is the purification of neurons of a particular type or stage of maturation from the differentiating culture as a whole[75] (Figure 1.2A). Isolation of a neural type of interest will also be invaluable for pinpointing the cell-autonomous effects of genetic variants. One proven strategy for neuronal isolation has been the production of “reporter” pluripotent stem cell lines using either transgenesis or gene targeting[76,77]. In general these methods have relied on using promoter sequences or other regulatory elements to delineate the expression of fluorescent proteins (GFP or RFP for example) or cell-surface antigens (E.G. H2kk, CD4, etc…) specifically to the neuronal or progenitor populations of interest[77,78]. When these transgenic cell lines are coupled to appropriate directed differentiation strategies they enable the isolation of the desired neuronal populations via fluorescent activated cell sorting (FACS)[77] or magnetic separation[78]. Purification of particular neurons allows for the establishment of a defined culture and facilitates longitudinal studies while also assuring the exclusion of progenitor cells thereby “time stamping” postmitotic neurons with a minimum *in vitro* age[79]. Additionally, such an approach should greatly facilitate gene expression studies by methods including RNA sequencing.

Given that expression of a particular gene, and by extension expression of reporter transgene harboring its regulatory elements, is often only sufficient to identify a broad population

of neural types[80], understanding the strengths and limitations of a given reporter will be critical for the proper design and interpretation of disease-related studies. For example, in Parkinson's disease it is well known that dopaminergic (DA) neurons degenerate during disease progression[81]. The generation of a human pluripotent stem cell line that reports on dopamine transporter (*DAT1*) expression would allow for the purification of DA neurons from heterogeneous neural cultures, allowing precise phenotypic comparisons between DA neurons of distinct phenotypes. However, in Parkinson's disease it is generally thought that only a subset of DA neurons in the midbrain are most susceptible to degeneration whereas other DA neurons both in the substantia nigra and other locations are more resistant[81]. Therefore failing to characterize the subtypes of DA neurons produced by a given protocol and collecting all *DAT1* expressing cells for analysis might obscure findings related to subtype-specific processes and lead to false generalizations about cellular events occurring in Parkinson's disease. The use of more specific markers for substantia nigra DA neurons which degenerate during Parkinson's disease progression, such as by using *PITX3*[82], or carefully characterizing the labeled *DAT1* expressing neurons produced during *in vitro* differentiation may, for instance, allow optimized systems for disease related discoveries to be established.

Although the expression of transcription factors can allow given neuronal types to be uniquely distinguished within certain regions of the brain, it is often the case that their transcription, or uniqueness of their expression is only limited to certain time-frames of development, or differentiation. Expression of the transcription factor HB9 can uniquely identify many motor neurons in the context of the developing spinal cord and neuronal cultures that have been caudalized to a spinal identity. However, *HB9* and reporter genes driven by its regulatory elements are silenced in many motor neurons as they mature[83]. Therefore, the timing of maximal utility for such reporters must also be considered. As a case in point, I have found that human ES cells reporting on motor neuron differentiation with an *Hb9*::GFP transgene are very useful for purifying immature neurons, but expected silencing of this reporter occurs in

motor neurons over time, suggesting it is of little use for monitoring their very long-term survival. As a result, depending on the tools at hand for monitoring the survival and long-term behavior of a given class of neurons, the identification of unique signatures of gene expression that persist through maturation of a given neuronal type, or the employment of genetic lineage tracing strategies, such as Cre/Lox are likely to be warranted.

Development of Disease Phenotypes in Culture

Another source of neuronal variation that could occlude disease-related phenotypes is disparities in the state of their functional maturation. Progenitors represent a particularly significant obstacle when neurons must be cultured for long periods of time in order to allow for maturation of phenotypes. This is because progenitors will continue to divide in culture as phenotypes develop over time, and can take over a post-mitotic neuronal culture given sufficient time. If this occurs, progenitor-related 'noise' can obscure any emerging 'signal' from disease-related phenotypes (Figure1. 2D). Furthermore, many of the properties that might be predicted to decline over time due to degenerative phenotypes, including soma size, neurite arborization, synapse number and aspects of physiological activity also change dramatically over the course of a neuron's maturation[84,85]. Thus if a given culture is composed of neurons with widely varying states of maturation, for instance through the action of lingering progenitors, then it may be difficult to distinguish young neurons that are not yet mature from older neurons which have begun to show signs of degeneration (Figure 1.2D). As I indicate above, purification of postmitotic neural types from progenitors using reporter genes to birthdate neuronal populations may help overcome this difficulty[72]. In instances where purification of a given neural type is not feasible, some of these difficulties may be overcome with pulse chase labeling techniques, including the incorporation of BrdU into dividing cells that can illuminate any irregularities in progenitor activity between cell lines and determine at which stage neurons were "born" into a culture[86]. Also, the use of anti-mitotics can reduce the production of immature neurons by

inhibiting progenitor cell proliferation. However, the neurons might suffer mild cytotoxic consequences[85]. Another possibility to reduce the contributions of progenitor cells to any phenotype is to use neurons generated from terminal cell types[87]. Recent experiments have demonstrated the possibility of reprogramming terminally differentiated cell types into post mitotic neurons relevant to many neurological disorders. These approaches seem to directly convert one terminal cell type into another, and therefore progenitor cell types are never introduced into the culture[87]. While these new techniques may one day allow for the production of homogeneous cultures of single terminal cell types, at present they generally result in a heterogeneous population consisting of the unconverted original cell type and partially reprogrammed cell types along with the cell type of interest, and thus suffer from similar heterogeneity-related difficulties encountered in cultures derived from directed neuronal differentiation protocols. Finally, it is important to remember that time in culture may not have a direct relationship with state of maturity of a given neuronal population[88]. If a functional state is required to monitor a phenotype in question, supplementing culture conditions with agents that promote the characteristic in question, for instance astrocyte conditioned media for the induction of synapse formation[89], may be essential.

Another question of significant importance, which has yet to be well addressed is: How long, should I expect to wait before a degenerative phenotype emerges? Certainly time to functional maturation may play a role in the schedule by which phenotypes might be expected to appear. However, there are also the more nebulous factors that influence the appearance of degenerative phenotypes *in vivo* that are associated with patient aging and are likely important contributors to many neurological conditions[90]. Currently, there is no clear relation that can be made between time in culture and “age” of a neuronal type *in vivo*. Some studies have demonstrated *in-vivo* transcriptional and epigenetic changes which occur during normal aging of the brain[91,92]. While this line of research is still in its infancy, it may one day be possible to use the expression of a few well-chosen genes or epigenetic alterations at specific loci to

benchmark *in-vitro* derived neurons to their *in-vivo* counterparts. When considering our current inability to quantitatively determine a neurons *in-vitro* age, again the action of progenitor cells could be detrimental to detecting any effects the age of a neuron may have on disease phenotypes, arguing once more that their activity should be monitored or eliminated (Figure 1.2B). Even given a more homogeneous culture of "aging" neurons, the length of time required to detect disease related changes is controversial and may vary substantially from case to case. For example, neuronal dysfunction and death resulting from chronic cellular stress induced by miss-folded or misprocessed proteins is a hypothesis relevant to many neurodegenerative diseases [93,94]. In order to study events that occur later in pathogenesis, such as the formation of large scale protein inclusions, it may be necessary to wait a sufficient length of time for mutant proteins to accumulate or additional poorly understood events to occur. For example, cytological accumulation of p62 punctate were not observed in neurons derived from Parkinson's patients after 30 days in culture[21]. However, after a total of seventy-five days in culture, the accumulation of p62 foci reminiscent to histological inclusions observed upon patient autopsy were found[21,95]. While these findings are encouraging, they also re-raise our question concerning the precise nature of changes that occur during long-term culture and the relation to aging, if any, to which neurons cultured for different durations of time correspond.

Another possibility is that many mutations predisposing a given neuronal type to degeneration may begin exerting their relevant negative effects as soon as they become expressed. For instance, immunocytochemistry has uncovered disease specific processing of mutant protein species in iPSC derived neurons harboring predisposing mutations to ALS and Alzheimer's Disease[96,97]. In cases where this is relevant, careful studies of the immediate changes induced in the most sensitive neuronal types may provide insight with therapeutic value, long before end-points of degeneration such as protein inclusion formation or neuron death are reached. Additionally, stressors that antagonize relevant neuronal defenses may also unmask or accelerate disease processes enabling their study in more manageable time

frames[22,98]. In fact such techniques have been successfully utilized in several studies to date[22,98]. Neurons derived from familial Parkinson's patients proved more sensitive to antagonists of proteasome activity, as well as inducers of oxidative and mitochondrial stress than controls[99]. While these results demonstrate the utility of stressors for magnifying disease associated phenotypes, there are some drawbacks to this approach. Many small molecules used as stressors are not wholly specific for the intended target and will modulate unintended pathways when used at higher concentrations[100]. The use of multiple small molecules or genetic approaches to impact the same pathway can validate a given targets importance to disease pathogenesis.

Modeling Neural Selectivity of Disease

Cell type specificity is a hallmark of many neurological disorders[101]. However, few attempts at revealing intrinsic differences between vulnerable and resistant neurons have been earnestly made[101]. The production and comparison of vulnerable and resistant neurons from pluripotent cell lines is an intriguing option for interrogating disease cell type selectivity (Figure 1.3B and C). For example, in ALS, spinal motor neurons are most sensitive to the toxic effects of mutant SOD1 protein[102]. However, ocular motor neurons that allow blinking and eye movement survive much longer in patients and mouse models with the disorder[103]. Comparing ocular and spinal motor neurons under various conditions (oxidative stress, neurotrophic factor withdraw, glutamate excitotoxicity) could reveal intrinsic susceptibilities in cellular pathways that underlie the hallmark cell type specificity of ALS. Additionally the use of multiple neuronal subtypes derived from large cohorts of patient specific pluripotent cell lines could reveal subtle variations in the identity of resistant and vulnerable neurons driven by distinct disease variants. As described above, such an approach could be useful to re-classify subgroups of patients previously aggregated based solely on patient diagnosis. The C9orf72 repeat expansion has been implicated in comorbid Frontal Temporal Degeneration (FTD), a disease which partially results from degeneration of von Economo neurons, whereas SOD1 has never been implicated

in FTD[5,104]. This divergence in ALS-causing mutations could naturally pave the way for experiments aimed at characterizing and differentiating the pathophysiology which leads to degeneration of motor vs. von Economo neurons. This could be achieved by using patient-derived neurons harboring different familial ALS mutations, including C9orf72 patient specific iPSCs which have comorbid FTD and SOD1 patient specific iPSCs, and characterizing their phenotypes in both motor and von Economo neurons. One obstacle to this type of research, however, is the limited availability of protocols for the differentiation of disease-resistant neurons. While significant effort has been devoted to the development of protocols to produce cells lost in disease, there has been less focus on producing cells resistant to disease processes. Faced with this challenge, it may be tempting to use the most convenient cell types for comparison based on available differentiation protocols. For example, protocols for the production of ocular motor neurons have not been established, whereas efficient methods for the generation of sensory neurons, another subtype unaffected during ALS disease progression, have been published[69]. Comparisons between spinal MNs and sensory neurons may provide insights into the subtype selectivity of ALS (Figure 1.3). However, selecting comparison cell types differing as little as possible from the cell type of interest, except in the vulnerability to the disease process, promises to provide the clearest path to understanding disease selectivity by eliminating many differences particular to cell fate but not relevant to the disease.

Figure 1.3

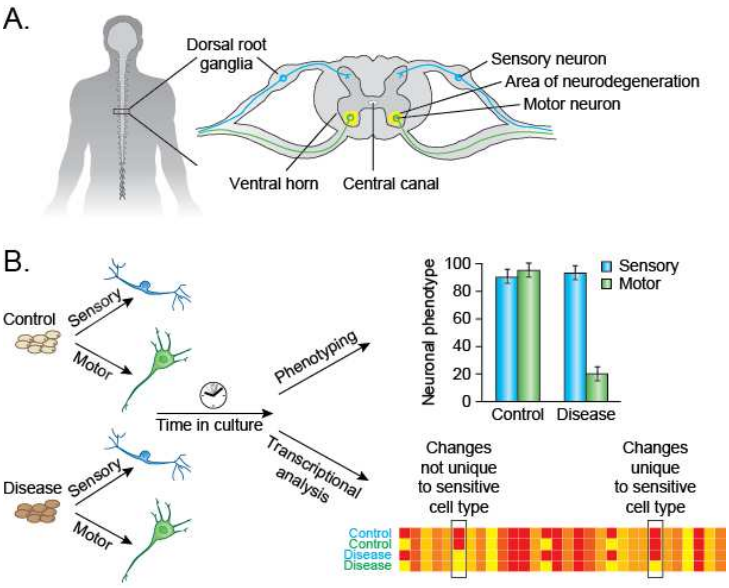

Figure 1.3 Observation of neurons resistant to the disease being studied could allow meaningful comparisons between vulnerable and resistant neurons derived from control and patient pluripotent cell lines. (A) During ALS disease progression spinal motor neurons are lost, resulting in paralysis and eventually death of the patient. Sensory neurons, however, remain viable. (B) In vitro comparisons between motor and sensory neurons derived from control and patient hiPSC lines may recapitulate the in vivo cell type selectivity of the disease. If reproducible phenotypes are uncovered, transcriptional analysis of the resistant and vulnerable neuronal populations may reveal mechanisms for the motor neuron specificity of ALS. (All graphs are hypothetical and not from actual data.)

Conclusion

Enthusiasm for the use of *in vitro* disease modeling has developed at a rapid pace. The exciting possibilities enabled by this technology have sparked many attempts in its application. Initial studies which utilized pluripotent stem cell technology to model disease pathology generally compared one or two patient derived pluripotent cell lines to a similarly sparse group of control cell lines. Results from these preliminary experiments may or may not stand the test of time given what I now know regarding variability amongst multiple pluripotent cell lines. As the black box of reprogramming and pluripotency has become illuminated, better practices in characterizing pluripotent cell lines have been established, allowing for quick elimination of atypical cell lines which could confound disease modeling results. Advances in reprogramming techniques have allowed for the generation of large cohorts of disease and control pluripotent

cell lines, allowing for the production of neurons from many individuals and thereby reducing the possibility of results being driven by aberrant cell lines. Furthermore, the rapid development of gene targeting techniques capable of modifying single base pairs in pluripotent stem cell lines now allows for comparison between isogenic stem cell derived neurons, emphasizing disease variant driven phenotypes. Clearly, the sophistication of disease modeling experiments has advanced a great deal. Screening large libraries of small molecules for their ability to alleviate phenotypes in neurons derived from patient specific iPSCs is being pursued by many labs today. A critical component to any screen is the robustness of the phenotype which any potential therapeutic must improve. The advances in disease modeling experiments described above will provide more reliable platforms for the development of large scale screens and are critical for the future of this line of research. Still, there is room for improvement in disease modeling experiments. Purification of the cell type implicated in disease pathology is not routine, and is often impossible when antigens specific to the degenerating neurons are unknown. Furthermore, the *in-vitro* maturation of neurons is still difficult to determine, making correlations with stages of patient disease pathology difficult. No doubt, as more defined disease specific molecular signatures are uncovered and *in-vitro* maturation of stem cell derived neurons becomes more predictable, better models capable of revealing previously unknown disease pathology will be established.

References

1 Andersen, P. M. & Al-Chalabi, A. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nature Reviews Neurology* **7**, 603-615 (2011).

2 Chiò, A. *et al.* Large proportion of amyotrophic lateral sclerosis cases in Sardinia due to a single founder mutation of the TARDBP gene. *Archives of neurology* **68**, 594-598 (2011).

3 van Es, M. A. *et al.* Large-scale SOD1 mutation screening provides evidence for genetic heterogeneity in amyotrophic lateral sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* **81**, 562-566 (2010).

4 Weber, M. *et al.* ALS patients with SOD1 mutations in Switzerland show very diverse phenotypes and extremely long survival. *Journal of Neurology, Neurosurgery & Psychiatry* **83**, 351-353 (2012).

5 Renton, A. E. *et al.* A Hexanucleotide Repeat Expansion in< i> C9ORF72</i> Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **72**, 257-268 (2011).

6 Forsberg, K. *et al.* Novel antibodies reveal inclusions containing non-native SOD1 in sporadic ALS patients. *PLoS ONE* **5**, e11552 (2010).

7 Bosco, D. *et al.* Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS. *Nature neuroscience* **13**, 1396-1403, doi:10.1038/nn.2660 (2010).

8 Andersen, P. M. Mutation in C9orf72 changes the boundaries of ALS and FTD. *Lancet Neurol* **11**, 205-207, doi:10.1016/s1474-4422(12)70020-0 (2012).

9 Gurney, M. E. *et al.* Motor neuron degeneration in mice that express a human Cu, Zn superoxide dismutase mutation. *SCIENCE-NEW YORK THEN WASHINGTON-*, 1772-1772 (1994).

10 Di Giorgio, F., Carrasco, M., Siao, M., Maniatis, T. & Eggan, K. Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model. *Nature neuroscience* **10**, 608-614, doi:10.1038/nn1885 (2007).

11 Gould, T. *et al.* Complete dissociation of motor neuron death from motor dysfunction by Bax deletion in a mouse model of ALS. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **26**, 8774-8786, doi:10.1523/jneurosci.2315-06.2006 (2006).

12 Deitch, J. S. *et al.* Phenotype of Transgenic Mice Carrying a Very Low Copy Number of the Mutant Human G93A Superoxide Dismutase-1 Gene Associated with Amyotrophic Lateral Sclerosis. *PLoS ONE* **9**, e99879 (2014).

13 Gladman, M., Cudkowicz, M. & Zinman, L. Enhancing clinical trials in neurodegenerative disorders: lessons from amyotrophic lateral sclerosis. *Curr Opin Neurol* **25**, 735-742, doi:10.1097/WCO.0b013e32835a309d (2012).

14 Bellin, M., Marchetto, M. C., Gage, F. H. & Mummery, C. L. Induced pluripotent stem cells: the new patient? *Nature Reviews Molecular Cell Biology* (2012).

15 Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872 (2007).

16 Dimos, J. *et al.* Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science (New York, N.Y.)* **321**, 1218-1221, doi:10.1126/science.1158799 (2008).

17 Park, I.-H. *et al.* Disease-specific induced pluripotent stem cells. *Cell* **134**, 877-886, doi:10.1016/j.cell.2008.07.041 (2008).

18 Liu, G.-H. *et al.* Progressive degeneration of human neural stem cells caused by pathogenic LRRK2. *Nature* **491**, 603-607, doi:10.1038/nature11557 (2012).

19 Han, S. S. W., Williams, L. A. & Eggan, K. C. Constructing and deconstructing stem cell models of neurological disease. *Neuron* **70**, 626-644 (2011).

20 Yagi, T. *et al.* Modeling familial Alzheimer's disease with induced pluripotent stem cells. *Human Molecular Genetics* **20**, 4530-4539 (2011).

21 Sánchez-Danés, A. *et al.* Disease-specific phenotypes in dopamine neurons from human iPS-based models of genetic and sporadic Parkinson's disease. *EMBO molecular medicine*, doi:10.1002/emmm.201200215 (2012).

22 Jeon, I. *et al.* Neuronal Properties, In Vivo Effects, and Pathology of a Huntington's Disease Patient-Derived Induced Pluripotent Stem Cells. *Stem cells* **30**, 2054-2062 (2012).

23 Ebert, A. *et al.* Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* **457**, 277-280, doi:10.1038/nature07677 (2009).

24 Mekhoubad, S. *et al.* Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* **10**, 595-609, doi:10.1016/j.stem.2012.02.014 (2012).

25 Marchetto, M. C. N. *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* **143**, 527-539 (2010).

26 Brennand, K. *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* **473**, 221-225, doi:10.1038/nature09915 (2011).

27 Boulting, G. L. *et al.* A functionally characterized test set of human induced pluripotent stem cells. *Nature biotechnology* **29**, 279-286 (2011).

28 Xu, C. *et al.* Feeder-free growth of undifferentiated human embryonic stem cells. *Nature biotechnology* **19**, 971-974 (2001).

29 Chen, A. E. *et al.* Optimal timing of inner cell mass isolation increases the efficiency of human embryonic stem cell derivation and allows generation of sibling cell lines. *Cell Stem Cell* **4**, 103 (2009).

30 Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-1147 (1998).

31 Di Giorgio, F. P., Boulting, G. L., Bobrowicz, S. & Eggan, K. C. Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. *Cell Stem Cell* **3**, 637-648, doi:S1934-5909(08)00522-5 [pii] 10.1016/j.stem.2008.09.017 [doi] (2008).

32 Marchetto, M. C. N. *et al.* Non-Cell-Autonomous Effect of Human SOD1< sup> G37R</sup> Astrocytes on Motor Neurons Derived from Human Embryonic Stem Cells. *Cell Stem Cell* **3**, 649-657 (2008).

33 Verlinsky, Y. *et al.* Human embryonic stem cell lines with genetic disorders. *Reproductive biomedicine online* **10**, 105-110 (2005).

34 Mateizel, I. *et al.* Derivation of human embryonic stem cell lines from embryos obtained after IVF and after PGD for monogenic disorders. *Human Reproduction* **21**, 503-511 (2006).

35 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, N.Y.)* **318**, 1917-1920, doi:10.1126/science.1151526 (2007).

36 Kiskinis, E. & Eggan, K. Progress toward the clinical application of patient-specific pluripotent stem cells. *J Clin Invest* **120**, 51-59, doi:10.1172/jci40553 (2010).

37 Chin, M. H. *et al.* Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* **5**, 111-123 (2009).

38 Müller, F.-J. *et al.* A bioinformatic assay for pluripotency in human cells. *Nature methods* **8**, 315-317 (2011).

39 Bock, C. *et al.* Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439-452 (2011).

40 Tomoda, K. *et al.* Derivation Conditions Impact X-Inactivation Status in Female Human Induced Pluripotent Stem Cells. *Cell Stem Cell* **11**, 91-99 (2012).

41 Ohi, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature cell biology* **13**, 541-549 (2011).

42 Daley, G. Q. *et al.* Broader implications of defining standards for the pluripotency of iPSCs. *Cell Stem Cell* **4**, 200-201; author reply 202, doi:S1934-5909(09)00064-2 [pii] 10.1016/j.stem.2009.02.009 [doi] (2009).

43 Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68-73 (2011).

44 Marchetto, M. C. N. *et al.* Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS One* **4**, e7076 (2009).

45 Polo, J. M. *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature biotechnology* **28**, 848-855 (2010).

46 Silva, S. S., Rowntree, R. K., Mekhoubad, S. & Lee, J. T. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proceedings of the National Academy of Sciences* **105**, 4820-4825 (2008).

47 Osafune, K. *et al.* Marked differences in differentiation propensity among human embryonic stem cell lines. *Nature biotechnology* **26**, 313-315 (2008).

48 Hu, B., Weick, J., Yu, J. & Ma, L. Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proceedings of the ...* (2010).

49 Kondo, T. *et al.* Modeling Alzheimer's disease with iPSCs reveals stress phenotypes associated with intracellular Aβ and differential drug responsiveness. *Cell Stem Cell* (2013).

50 Consortium, H. *et al.* Induced Pluripotent Stem Cells from Patients with Huntington's Disease Show CAG-Repeat-Expansion-Associated Phenotypes. *Cell Stem Cell* **11**, 264-278 (2012).

51 Chang, T. *et al.* Brief Report: Phenotypic Rescue of Induced Pluripotent Stem Cell-Derived Motoneurons of a Spinal Muscular Atrophy Patient. *Stem cells* **29**, 2090-2093 (2011).

52 Soldner, F. *et al.* Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* **146**, 318-331, doi:10.1016/j.cell.2011.06.019 (2011).

53 Kiskinis, E. *et al.* Pathways Disrupted in Human ALS Motor Neurons Identified through Genetic Correction of Mutant SOD1. *Cell Stem Cell*, doi:10.1016/j.stem.2014.03.004 (2014).

54 An, M. C. *et al.* Genetic correction of Huntington's disease phenotypes in induced pluripotent stem cells. *Cell Stem Cell* (2012).

55 Khan, I. F., Hirata, R. K. & Russell, D. W. AAV-mediated gene targeting methods for human cells. *Nature protocols* **6**, 482-501 (2011).

56 Aizawa, E. *et al.* Efficient and Accurate Homologous Recombination in hESCs and hiPSCs Using Helper-dependent Adenoviral Vectors. *Molecular therapy : the journal of the American Society of Gene Therapy* **20**, 424-431, doi:10.1038/mt.2011.266 (2012).

57 Andersen, M. S., Sørensen, C. B., Bolund, L. & Jensen, T. G. Mechanisms underlying targeted gene correction using chimeric RNA/DNA and single-stranded DNA oligonucleotides. *Journal of molecular medicine* **80**, 770-781 (2002).

58 Zou, C. *et al.* Efficient Derivation and Genetic Modifications of Human Pluripotent Stem Cells on Engineered Human Feeder Cell Lines. *Stem cells and development*, doi:10.1089/scd.2011.0688 (2012).

59 Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* (2013).

60 Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics* **11**, 636-646 (2010).

61 Reinhardt, P. *et al.* Genetic Correction of a LRRK2 Mutation in Human iPSCs Links Parkinsonian Neurodegeneration to ERK-Dependent Changes in Gene Expression. *Cell Stem Cell* **12**, 354-367 (2013).

62 Corti, S. *et al.* Genetic Correction of Human Induced Pluripotent Stem Cells from Patients with Spinal Muscular Atrophy. *Science Translational Medicine* **4**, 165ra162-165ra162 (2012).

63 Lombardo, A. *et al.* Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nature methods* **8**, 861-869, doi:10.1038/nmeth.1674 (2011).

64 Qiurong, D. *et al.* A TALEN Genome-Editing System for Generating Human Stem Cell-Based Disease Models. *Cell Stem Cell*, doi:10.1016/j.stem.2012.11.011 (2012).

65 Veres, A. *et al.* Low Incidence of Off-Target Mutations in Individual CRISPR-Cas9 and TALEN Targeted Human Stem Cell Clones Detected by Whole-Genome Sequencing. *Cell Stem Cell* **15**, 27-30 (2014).

66 Murry, C. E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661-680, doi:S0092-8674(08)00216-X [pii] 10.1016/j.cell.2008.02.008 [doi] (2008).

67 Wichterle, H., Lieberam, I. & Porter, J. Directed differentiation of embryonic stem cells into motor neurons. *Cell* (2002).

68 Kanning, K. C., Kaplan, A. & Henderson, C. E. Motor neuron diversity in development and disease. *Annu Rev Neurosci* **33**, 409-440, doi:10.1146/annurev.neuro.051508.135722 (2010).

69 Chambers, S. *et al.* Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nature biotechnology* **30**, 715-720, doi:10.1038/nbt.2249 (2012).

70 Wataya, T. *et al.* Minimization of exogenous signals in ES cell culture induces rostral hypothalamic differentiation. *Proceedings of the National Academy of Sciences* **105**, 11796-11801 (2008).

71 Kriks, S. *et al.* Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson/'s disease. *Nature* (2011).

72 Yuan, S. *et al.* Cell-surface marker signatures for the isolation of neural stem cells, glia and neurons derived from human pluripotent stem cells. *PLoS One* **6**, doi:10.1371/journal.pone.0017540 (2011).

73 Lobsiger, C. S. & Cleveland, D. W. Glial cells as intrinsic components of non-cell-autonomous neurodegenerative disease. *Nat Neurosci* **10**, 1355-1360 (2007).

74 Amoroso, M. W. *et al.* Accelerated High-Yield Generation of Limb-Innervating Motor Neurons from Human Stem Cells. *J Neurosci* **33**, 574-586, doi:10.1523/jneurosci.0906-12.2013 (2013).

75 Sergent-Tanguy, S., Chagneau, C., Neveu, I. & Naveilhan, P. Fluorescent activated cell sorting (FACS): a rapid and reliable method to estimate the number of neurons in a mixed population. *Journal of neuroscience methods* **129**, 73-79 (2003).

76 Hockemeyer, D. *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nature biotechnology* **27**, 851-857, doi:10.1038/nbt.1562 (2009).

77 Singh Roy, N. *et al.* Enhancer-specified GFP-based FACS purification of human spinal motor neurons from embryonic stem cells. *Experimental neurology* **196**, 224-234 (2005).

78 Vodyanik, M. A., Thomson, J. A. & Slukvin, I. I. Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. *Blood* **108**, 2095-2105 (2006).

79 Bilican, B. *et al.* Mutant induced pluripotent stem cell lines recapitulate aspects of TDP-43 proteinopathies and reveal cell-specific vulnerability. *Proceedings of the National Academy of Sciences* **109**, 5803-5808 (2012).

80 Molyneaux, B. J., Arlotta, P., Menezes, J. R. L. & Macklis, J. D. Neuronal subtype specification in the cerebral cortex. *Nature Reviews Neuroscience* **8**, 427-437 (2007).

81 Dawson, T. M. & Dawson, V. L. Neuroprotective and neurorestorative strategies for Parkinson's disease. *Nature neuroscience* **5**, 1058-1061 (2002).

82 Nunes, I., Tovmasian, L. T., Silva, R. M., Burke, R. E. & Goff, S. P. Pitx3 is required for development of substantia nigra dopaminergic neurons. *Proceedings of the National Academy of Sciences* **100**, 4245-4250 (2003).

83 von Steyern, F. V. *et al.* The homeodomain transcription factors Islet 1 and HB9 are expressed in adult alpha and gamma motoneurons identified by selective retrograde tracing. *European Journal of Neuroscience* **11**, 2093-2102 (2008).

84 Takazawa, T. *et al.* Maturation of Spinal Motor Neurons Derived from Human Embryonic Stem Cells. *PLoS One* **7**, e40154 (2012).

85 Kim, J.-E. *et al.* Investigating synapse formation and function using human pluripotent stem cell-derived neurons. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 3005-3010, doi:10.1073/pnas.1007753108 (2011).

86 Juurlink, B. H. J. & Walz, W. Neural cell culture techniques. *Neuromethods* **33**, 53-102 (1998).

87 Pang, Z. P. *et al.* Induction of human neuronal cells by defined transcription factors. *Nature* **476**, 220-223 (2011).

88 Hester, M. *et al.* Rapid and efficient generation of functional motor neurons from human pluripotent stem cells using gene delivered transcription factor codes. *Molecular therapy : the journal of the American Society of Gene Therapy* **19**, 1905-1912, doi:10.1038/mt.2011.135 (2011).

89 Christopherson, K. S. *et al.* Thrombospondins are astrocyte-secreted proteins that promote CNS synaptogenesis. *Cell* **120**, 421-433 (2005).

90 Keller, J., Huang, F. & Markesbery, W. Decreased levels of proteasome activity and proteasome expression in aging spinal cord. *Neuroscience* **98**, 149-156 (2000).

91 Erraji-Benchekroun, L. *et al.* Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. *Biological psychiatry* **57**, 549-558 (2005).

92 Cheung, I. *et al.* Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. *Proceedings of the National Academy of Sciences* **107**, 8824-8829 (2010).

93 Wong, E. & Cuervo, A. M. Autophagy gone awry in neurodegenerative diseases. *Nature neuroscience* **13**, 805-811 (2010).

94 Matus, S., Glimcher, L. H. & Hetz, C. Protein folding stress in neurodegenerative diseases: a glimpse into the ER. *Current opinion in cell biology* **23**, 239-252 (2011).

95 Alegre-Abarrategui, J., Ansorge, O., Esiri, M. & Wade-Martins, R. LRRK2 is a component of granular alpha-synuclein pathology in the brainstem of Parkinson's disease. *Neuropathology and applied neurobiology* **34**, 272-283 (2007).

96 Egawa, N. *et al.* Drug screening for ALS using patient-specific induced pluripotent stem cells. *Sci Transl Med* **4**, 145ra104, doi:10.1126/scitranslmed.3004052 (2012).

97 Shi, Y. *et al.* A human stem cell model of early Alzheimer's disease pathology in Down syndrome. *Science Translational Medicine* **4**, 124ra129-124ra129 (2012).

98 Koch, P. *et al.* Excitation-induced ataxin-3 aggregation in neurons from patients with Machado-Joseph disease. *Nature* (2011).

99 Nguyen, H. N. *et al.* LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. *Cell Stem Cell* **8**, 267-280 (2011).

100 Fabian, M. A. *et al.* A small molecule–kinase interaction map for clinical kinase inhibitors. *Nature biotechnology* **23**, 329-336 (2005).

101 Saxena, S., Cabuy, E. & Caroni, P. A role for motoneuron subtype–selective ER stress in disease manifestations of FALS mice. *Nature neuroscience* **12**, 627-636 (2009).

102 Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59-62 (1993).

103 Brockington, A. *et al.* Unravelling the enigma of selective vulnerability in neurodegeneration: motor neurons resistant to degeneration in ALS show distinct gene

expression characteristics and decreased susceptibility to excitotoxicity. *Acta neuropathologica*, 1-15 (2013).

104 Seeley, W. W. *et al.* Early frontotemporal dementia targets neurons unique to apes and humans. *Annals of neurology* **60**, 660-667 (2006).

Chapter 2: Pathways Disrupted in Human ALS Motor Neurons Identified Through Genetic Correction of Mutant *SOD1*

The majority of this chapter was previously published in

Kiskinis E*, **Sandoe J***, Williams LA, Boulting GL, Moccia R, Wainger BJ, Han S, Peng T, Thams T, Mikkilineni S, Mellin C, Merkle FT, Ziller M, Oakley D, Ichida J, Dicostanza S, Atwater N, Maeder ML, Goodwin MJ, Joung JK, Woolf CJ, Brown RH, Eggan KC. Pathways Disrupted in Human ALS Motor Neurons Identified through Genetic Correction of Mutant SOD1. *Cell Stem Cell* (2014), http://dx.doi.org/10.1016/j.stem.2014.03.004

Wainger BJ*, Kiskinis E*, Mellin C, Wiskow O, Han S, **Sandoe J**, Perez NP, Williams LA, Lee S, Boulting GL, Berry JD, Brown RH, Cudkowicz ME, Bean BP, Eggan KC & Woolf CJ. Intrinsic Membrane Hyperexcitability Contributes to Death of ALS Patient-Derived Motor Neurons. *Cell Reports* (2014), http://dx.doi.org/10.1016/j.celrep.2014.03.019

Introduction

ALS is a fatal neurological condition characterized by selective death of motor neurons (MNs) [1,2]. Both classical linkage studies and modern genomic approaches involving DNA sequencing have demonstrated that ALS can be caused by a variety of mutations in more than two dozen genes acting on diverse cellular functions[3,4]. Mutations in *superoxide dismutase 1* (*SOD1*)[5] were originally identified through their autosomal dominant inheritance pattern. More recently, genome wide association studies, DNA sequencing efforts and linkage analysis have all contributed to the identification of a hexanucleotide repeat expansion at *C9orf72* as a likely cause of ALS in a substantial fraction of both familial and sporadic cases [6].

Importantly, the discovery of *SOD1* mutations led to a widely studied transgenic mouse model of ALS[7]. While indisputably valuable, these animals as well as many cell-based models, overexpress heterologous human SOD1 at super-physiological levels[7-9]. Therefore, it is generally accepted that findings from these animals carry the caveat that they could be artifacts of protein overexpression[10,11]. Furthermore, there is little information on how SOD1 impacts human MNs, leaving open to what extent results from the *SOD1* mouse model are of relevance to understanding disease in the human nervous system.

In addition, identification of patient mutations in other genes, such as *TDP43*[12] and *C9orf72*[6] have not yet translated into the creation of animal models that are as widely accepted as *SOD1* transgenic mice[3,13]. As a result, detailed understanding of whether the many changes in MNs induced by overexpression of mutant, human *SOD1* are also relevant to those forms of ALS caused by mutations in other loci, has been slow to develop. Investigations of whether there are shared mechanisms of MN disease are essential as they could inform the selection of pathways for therapeutic intervention with the greatest relevance to a broader patient population.

We and others have proposed that induced pluripotent stem cells (iPSCs) from ALS patients[14-16] and their differentiation into spinal[17,18] could complement existing animal models,

allowing hypotheses to be tested in human MNs expressing mutations in a variety of ALS causing loci, all at physiological levels[19-21].

Here, we have combined reprogramming and stem cell differentiation approaches with genome engineering and RNA sequencing technologies to identify the transcriptional and functional changes induced by the *SOD1A4V* mutation in human MNs. In addition to supporting hypotheses concerning the actions of mutant SOD1 protein developed using transgenic mouse models, such as the disruption of mitochondrial function and transport, our studies identified novel mechanisms that may contribute to MN dysfunction. Notably, we found that mutant *SOD1* disrupts a delicate balance between ER stress and neuronal excitability that is inherent to MNs. Finally, studies using iPSCs derived from patients harboring *C9orf72* repeat expansions indicate that at least a subset of the changes induced by mutant *SOD1* in human MNs are relevant to both forms of ALS.

Results

Generation of iPSCs and Functional Motor Neurons from *SOD1$^{+/A4V}$* ALS Patients.

We derived skin fibroblasts from two female ALS patients (study participants 39 and RB9) carrying the same dominantly acting *SOD1*A4V mutation (*SOD1$^{+/A4V}$*), then generated iPSCs via retroviral transduction of *OCT4*, *SOX2* and *KLF4* (Figure 2.1). Cell lines stably exhibited typical human pluripotent stem cell morphology with large nuclear-to-cytoplasmic ratios and prominent nucleoli, growing in dense colonies with well-defined borders, and immuno-positive for human pluripotency markers including nuclear NANOG and the cell surface antigen TRA-1-81 (Figure 2.1B-C). The pluripotent profile and the differentiation potential of the iPSC lines was also verified using a lineage scorecard assay (Figure 2.1D-E) (Bock et al., 2011).

Figure 2.1

A

Patient	Diagnosis	Genotype	Age of onset	Sex	Age at biopsy	Reprog. factors
39	fALS	*SOD1*$^{+/A4V}$	43y	F	44	O,S,K
RB9	fALS	*SOD1*$^{+/A4V}$	45y	F	50	O,S,K
11	Healthy	*SOD1*$^{+/+}$	N.A.	M	37	O,S,K
18	Healthy	*SOD1*$^{+/+}$	N.A.	F	48	O,S,K

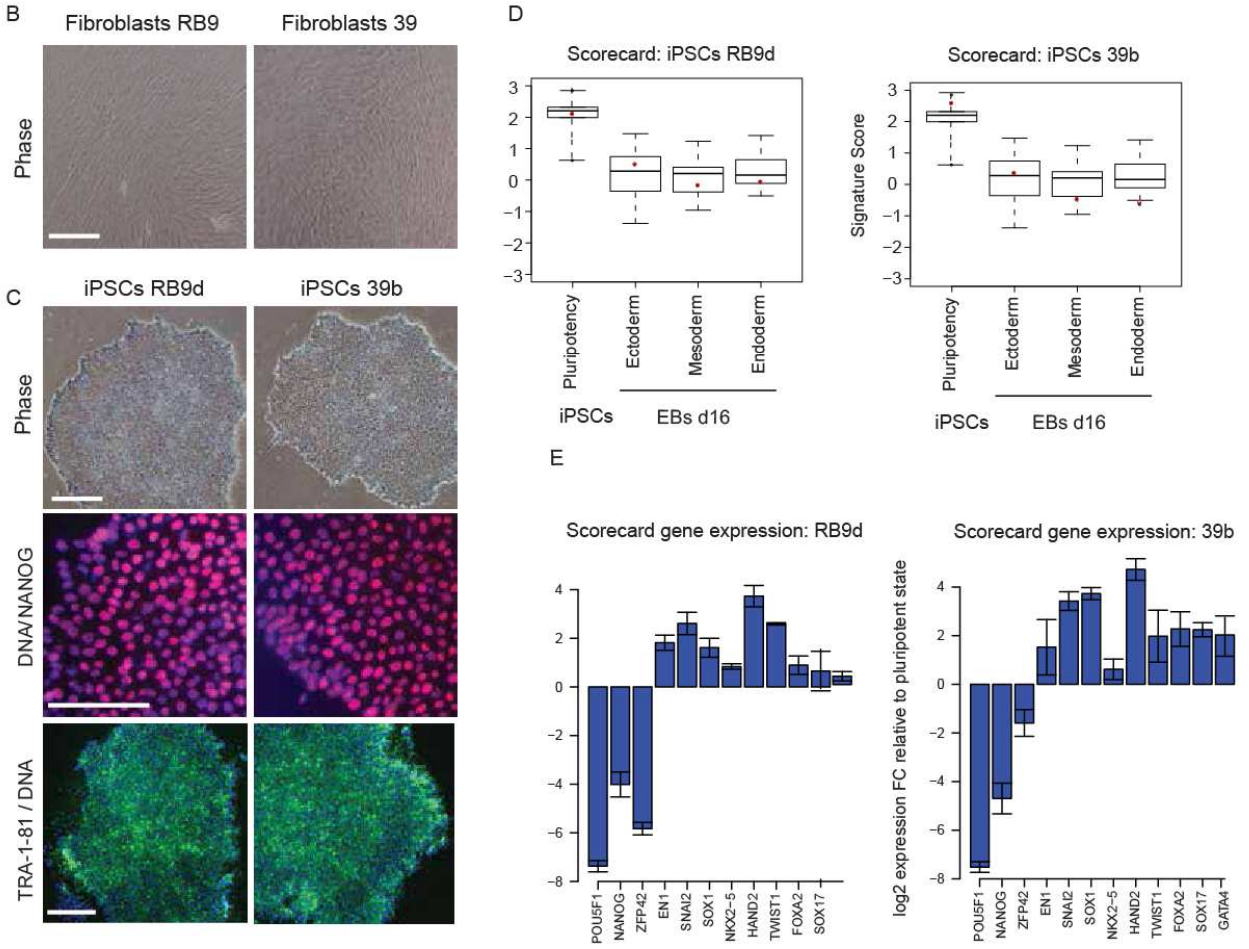

Figure 2.1 Generation and Characterization of Patient-Specific iPS Cells. (A) Summary of patients selected for this study. fALS, familial ALS, O: OCT4, S: SOX2, K: KLF4, N.A: not applicable. (B) Fibroblasts derived from skin biopsies of ALS-SOD1+/A4V patients. (C) iPSC colonies are morphologically identical to hESC colonies, and express NANOG and TRA-1-80, unlike the patient fibroblasts from which they were derived (scale bar=100µm). (D) Scorecard analysis of pluripotency genes for iPSCs and for three germ layer differentiation after 16 days, the red dots indicate values for each of the two lines. (E) Gene expression for pluripotency genes, ectoderm, mesoderm and endoderm genes after 16 days of undirected differentiation.

To produce MNs, we employed a 24-day differentiation protocol (Figure 2.2A) based on neuralization by dual SMAD inhibition[22] and MN specification through exposure to retinoic acid and a small molecule agonist of sonic hedgehog signaling [17,22]. *SOD1*$^{+/A4V}$ iPSC lines generated MNs that expressed the neuronal-specific class III β-TUBULIN (TUJ1) as well as the MN-

specific transcription factors ISLET 1/2 (ISL) and HB9 (Figure 2.2B)[23]. These MNs also expressed microtubule-associated protein 2 (MAP2) and choline acetyltransferase (ChAT) (Figure 2.2B)[24]. This differentiation strategy generated populations in which >97% of all cells were TUJ1+ neurons. Of these neurons, 21-38% were ISL+ MNs (n=6) (Figure 2.2C). To assess their functionality, we transduced d15 MNs with a lentivirus encoding RFP under control of the *Hb9* promoter (*Hb9::*RFP)[25] and performed whole-cell patch clamp recordings. RFP+ MNs possessed sodium and potassium currents, fired action potentials, and responded to excitatory and inhibitory transmitters as previously shown for other iPSC-derived MNs (Figure 2.2F-G)[17].

Figure 2.2

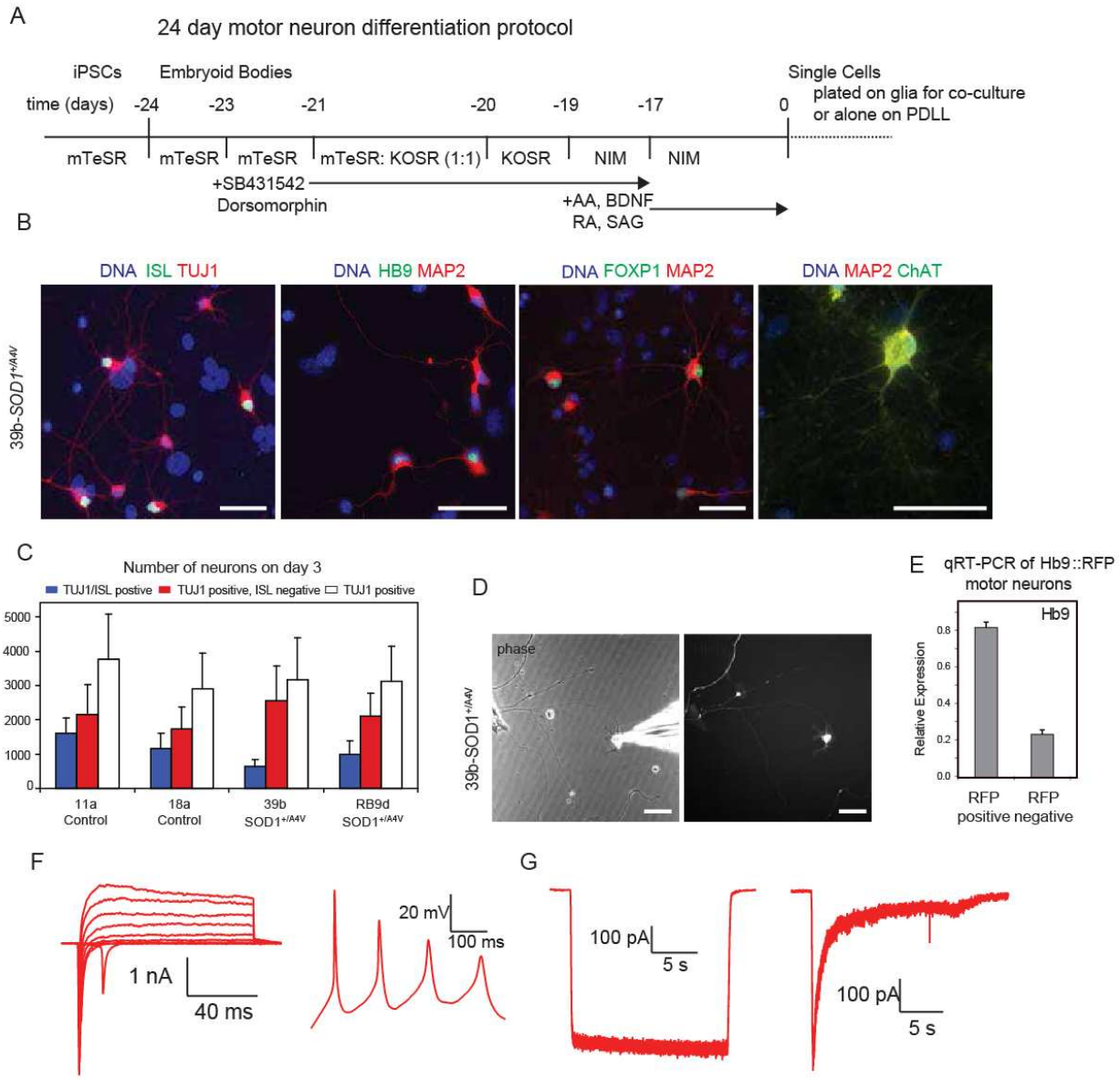

Figure 2.2 Differentiation and Characterization of Patient-Specific iPS cell-derived Motor Neurons. (A) Schematic of differentiation protocol (see methods for details). (B-C) Differentiations typically yielded neurons that expressed TUJ1, the MN transcription factors ISL and HB9, as well as MAP2. Many ISL+ MNs were FOXP1high+, at 30 days post-dissociation, ISL+ MNs were ChAT positive. (D-G) Hb9::RFP positive MNs are electrophysiologicaly active. RFP+ cells show a 6-fold enrichment in Hb9 expression levels than the RFP negative fraction. In whole-cell patch clamping recordings of day 15 neurons in co-culture with glia, MNs exhibited initial inward sodium currents followed by outward potassium currents upon depolarization in voltage-clamp mode, as well as repetitive action potential firing in current-clamp mode, as well as kainate and GABA responses (all scale bars=50µm).

Increased Apoptosis and Altered Morphometry in *SOD1$^{+/A4V}$* Motor Neurons.

Having demonstrated that we could produce MNs from *SOD1$^{+/A4V}$* patients, we proceeded to ask

whether these neurons might manifest a phenotype distinct from controls under standard culture conditions. To address this, we compared *SOD1*$^{+/A4V}$ MNs with MNs produced in parallel from two control iPSC lines (11a, 18a)[17], selected based on their similarity in neuronal differentiation capacity, iPSC reprogramming method and donor age. Differentiated preparations were plated on primary glial monolayers and the total number of ISL/TUJ1+ MNs was assessed after 3 and 30 days in culture (Figure 2.3A). We found that in comparison to the number of MNs present in cultures made from each line at day 3, there were significantly fewer *SOD1*$^{+/A4V}$ MNs at day 30 ($n=3$, $m>8000$, $P<0.05$)(Figure 2.3B).

Interestingly, the decline in *SOD1*$^{+/A4V}$ cultures seemed to be specific to MNs, rather than reflective of an overall deficit in neurons; with no significant difference in the number of ISL-negative, TUJ1+, presumptive non-MNs between cases and controls at day 30 ($n=3$, $m>25,000$)(Figure 2.3C).

We considered two explanations for the selective decline in MN number in *SOD1*$^{+/A4V}$ cultures. We first reasoned that if MN progenitor cells remained in the cultures, and if *SOD1*$^{+/A4V}$ progenitors were less abundant or functional, then the *SOD1*$^{+/A4V}$ MN number might selectively decline over time. To test this, we monitored progenitor activity via long-term BrdU incorporation (Figure 2.3D-E, G). When we chronically administered BrdU to cultures from day 0 to day 30, then assessed BrdU incorporation, we observed that only 2-3% of MNs were labeled (BrdU+, ISL+, TUJ1+) and that this modest rate of labeling was similar in control and *SOD1*$^{+/A4V}$ cultures (Figure 2.3D-E). These findings suggest that the vast majority (>97%) of MNs in our cultures were postmitotic prior to “day 0” and that the selective decline of MNs we observed in *SOD1*$^{+/A4V}$ cultures was not due to a decrease in progenitor activity.

The second hypothesis was that the selective decline of MN number in *SOD1*$^{+/A4V}$ cultures might result from an increased rate of cell death. We therefore used the BrdU incorporation strategy to specifically quantify only the MNs born before day 0. Using this criterion, we again found that significantly more *SOD1*$^{+/A4V}$ MNs were lost over the 30 days in

culture (m=3,686, P<0.05)(Figure 2.3F). To determine whether increased apoptosis caused the preferential loss of postmitotic *SOD1*$^{+/A4V}$ MNs, we performed TUNEL staining (Figure 2.3G). At day 21 of culture we found an increase in TUNEL+ cells in the two *SOD1*$^{+/A4V}$ cases compared to the two controls (n=2, P<0.05).

Figure 2.3

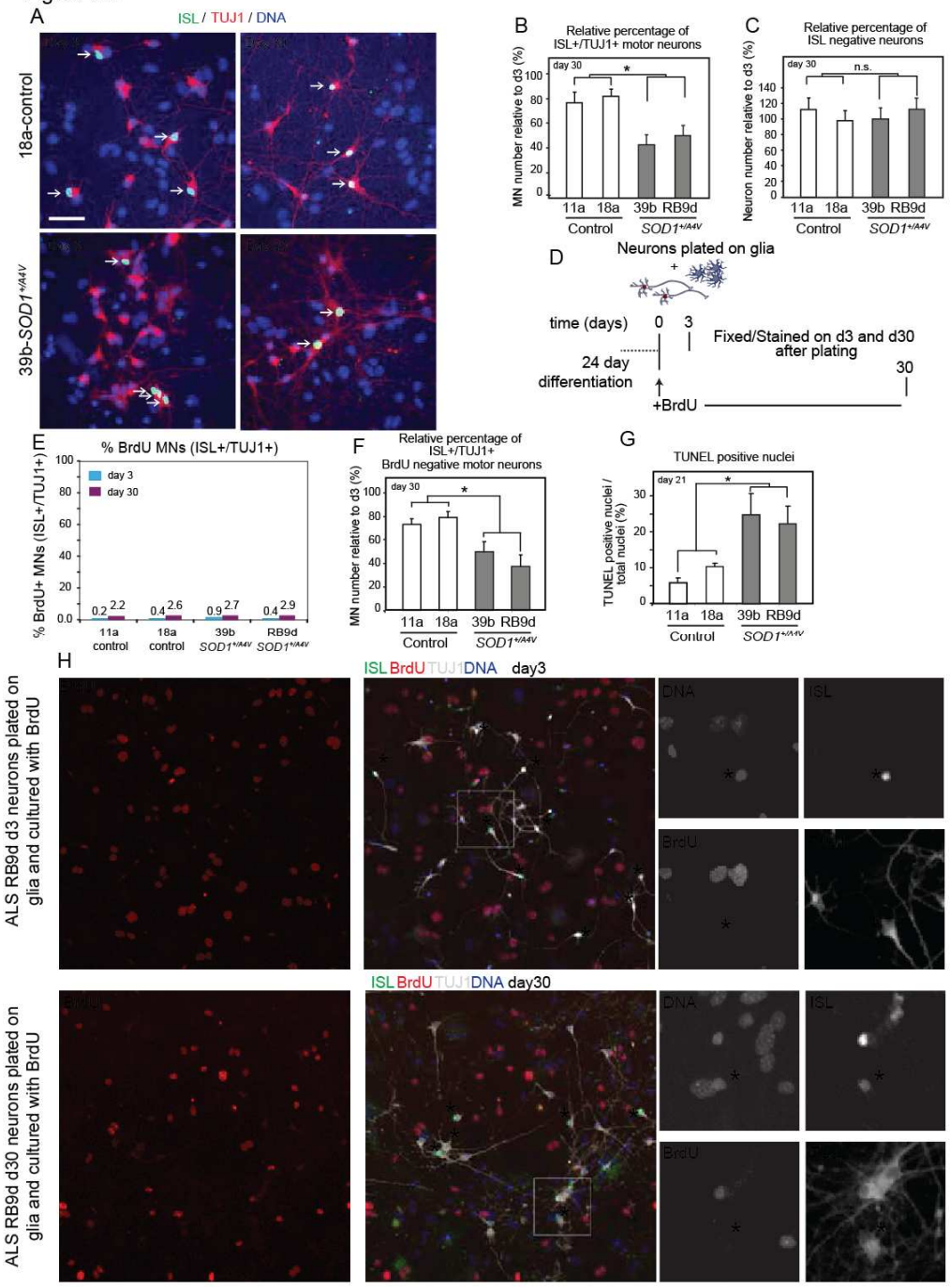

Figure 2.3 (Continued) iPSC-Derived MNs from SOD1+/A4V ALS Patients Exhibit Survival Differences Relative to Healthy Controls (A) Neuronal cultures on glial monolayers 3 and 30 days postdifferentiation from control and SOD1+/A4V iPSCs. MNs costained for ISL and TUJ1 are indicated by white arrows (scale bar = 50 mm). (B and C) Quantifications of ISL+ MNs (n = 3, m > 8,000, ±SEM, p < 0.05) (B) and ISL- neurons (n = 3, m > 25,000, ±SEM, p < 0.05) (C) after 30 days in culture. (D-E) Schematic of experimental strategy to evaluate the level of MN proliferation after dissociation; 1μM BrdU was added in plated co-cultures from days 0-30 and percentage of positive cells was analyzed. Between 0.2-0.8% of all MNs incorporate BrdU between days 0-3 while 2-2.8% between days 0-30 across all the iPSC lines. Importantly the mitoticly active supporting glial cells were exposed to BrdU in this experiment and many are BrdU+ (large DAPI nuclei). (H) A repre-sentative image of BrdU incorporation after the addition of 1μM BrdU to dissociated MN cultures after 3 and 30 days. Top right inserts exhibit an ISL+/BrdU negative MN, bottom inserts exhibit two ISL+ MNs, with only one being postive for BrdU also. Single arrows indicates BrdU+/ISL+ MNs; Double arrows indicate BrdU+/ISL negative cells; Asterisks indicate BrdU negative/ISL+ MNs, (scale bar=30μm). (E) Quantifications of MN numbers that are BrdU- after 30 days in culture (n = 1, m > 3,600, ±SD, p < 0.05). (F)Differential motor neurogenesis does not explain the lower numbers in SOD1+/A4V cases. (G) Quantifications of TUNEL+ nuclei of neuronal cultures without glia after 21 days in culture (n = 2, m > 13,000, ±SD, p < 0.05).

We also found that the pro-apoptotic phenotype we identified in *SOD1*$^{+/A4V}$ human MNs was accompanied by altered morphological characteristics similar to those seen in patients and in the *SOD1*^{G93A} mouse model[26,27]. In particular, 30-day old *SOD1*$^{+/A4V}$ MNs exhibited a significant reduction in relative soma size and fewer, as well as shorter processes compared to controls (n=3, m=88, P<0.01) (Figure 2.4A-H). Again, these effects were not observed in ISL-negative, TUJ1+ non-MNs (n=3, m=90) (Figure 2.4D). Interestingly, further analysis suggested that larger MNs were most vulnerable, as MNs larger than 150μm^2 made up less than 40% of the total MN population in *SOD1*$^{+/A4V}$ cases, compared to ~65% in the case of controls (Figure 2.4E). Because we observed variability with regard to baseline soma size between experiments in these studies, we fit the cell size data with a linear regression model. We did find that experiment number (that is, which differentiation) was an independent, significant predictor of absolute cell size ($p=2.8\times10^{-9}$). However, despite this, the difference between SOD1A4V cases and controls was still highly significant (35.2 μm^2 (SD 5.4, $p=2.0\times10^{-10}$). Thus in short, the absolute base-line size of neurons changed from experiment to experiment, however, the relative size distinction was always conserved with SOD1A4V motor neurons proving to be smaller than controls. Therefore, we presented the size of the ALS motor neurons normalized to that of control motor neurons from the same differentiation.

Figure 2.4

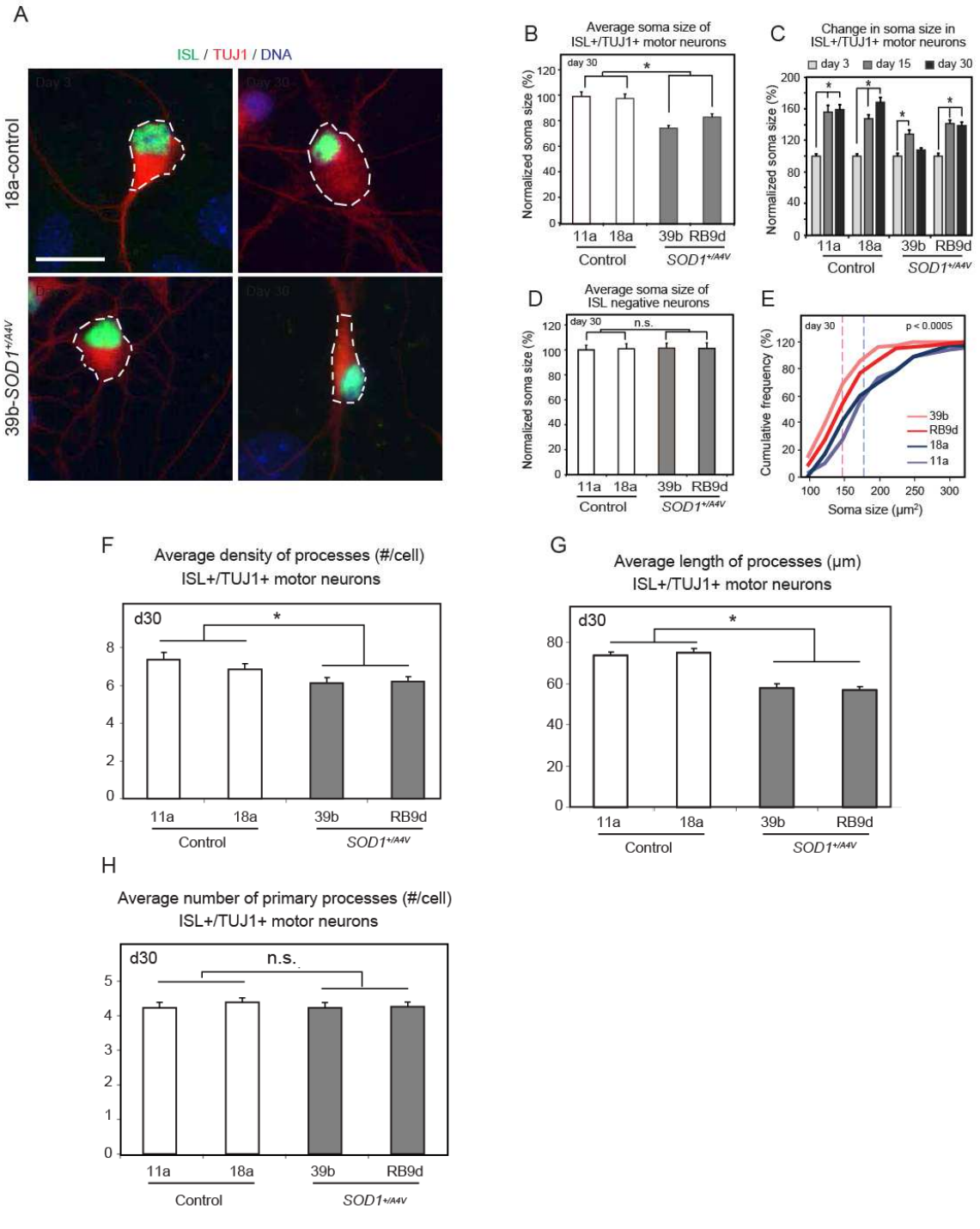
A
ISL / TUJ1 / DNA
18a-control
39b-SOD1+/A4V
B
Average soma size of ISL+/TUJ1+ motor neurons
day 30
Normalized soma size (%)
11a 18a 39b RB9d
Control
SOD1+/A4V
C
Change in soma size in ISL+/TUJ1+ motor neurons
day 3 day 15 day 30
Normalized soma size (%)
11a 18a 39b RB9d
Control
SOD1+/A4V
D
Average soma size of ISL negative neurons
day 30
n.s.
Normalized soma size (%)
11a 18a 39b RB9d
Control
SOD1+/A4V
E
day 30
p < 0.0005
Cumulative frequency (%)
39b
RB9d
18a
11a
Soma size (μm2)
F
Average density of processes (#/cell)
ISL+/TUJ1+ motor neurons
d30
11a 18a 39b RB9d
Control
SOD1+/A4V
G
Average length of processes (μm)
ISL+/TUJ1+ motor neurons
d30
11a 18a 39b RB9d
Control
SOD1+/A4V
H
Average number of primary processes (#/cell)
ISL+/TUJ1+ motor neurons
d30
n.s.
11a 18a 39b RB9d
Control
SOD1+/A4V

Figure 2.4 (Continued) iPSC-Derived MNs from SOD1+/A4V ALS Patients Exhibit Morphometric Differences Relative to Healthy Controls (A) Representative images of measured soma size (white-dotted circumference) of control and SOD1+/A4V MNs (scale bar = 20 mm). (B) Quantifications of ISL+ MN soma size with values normalized to those of controls (n = 3, m = 340, ±SEM, $p < 0.01$). (C) MN soma size after 3, 15, and 30 days in culture normalized to day 3 for each cell line. Although MNs increase in size over time in all four cell lines, they do so to a lesser degree in SOD1+/A4V cases. (D) Quantifications of ISL-negative neuron soma size with values normalized to those of controls (n = 3, m = 446, ±SEM, p = n.s.). (E) Cumulative frequency graphs of MN soma size after 30 days in culture. Dotted lines indicate averages for control and disease. n.s., not significant; n = experiment; m = cell number. (F) Quantifications of MN processes, evaluated as total number of TUJ1-positive processes per cell showed that SOD1+/A4V MNs are less elaborate after 30 days in culture; (G) and have shorter processes (n=3, $P < 0.05$, ± s.e.m.), (H) but not fewer primary processes.

To gain a sense of whether the survival difference we observed between *SOD1* mutant and control motor neurons was cell autonomous, we cultured preparations of motor neurons on primary glia from the $SOD1^{G93A}$ mouse model. In this context, and as expected based on previous studies[28], we found that both control and patient derived motor neurons survived significantly more poorly on $SOD1^{G93A}$ glia than on control glia. However, $SOD1^{+/A4V}$ MNs survived even more poorly than control MNs on $SOD1^{G93A}$ glia (Figure 2.5B). Thus in both cases the iPSC-derived MNs we utilized were sensitive to well-studied models of non-cell autonomous toxicity involving mutant SOD1. However, $SOD1^{+/A4V}$ MNs showed an additional survival deficit beyond that induced by the mutant glia, suggesting that at least in part, there was a cell-autonomous contribution to the effects described above (Figure 2.3) with $SOD1^{+/A4V}$ MNs on control glia.

Figure 2.5

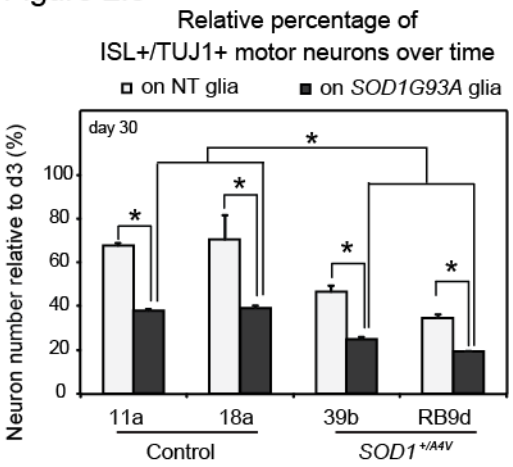

Figure 2.5. MNs are sensitive to the toxic effects of primary glia from SOD1G93A mice.

Gene Targeting and Correction of the *SOD1*A4V Mutation. Distinct iPSC lines can harbor significant variation in transcription and DNA methylation[29]. This variation, in conjunction with a cell line's particular genetic make-up, give it its own distinctive functional characteristics [17], which may confound disease modeling experiments[30]. Additionally, ALS patients can harbor more than one disease-causing genetic variant[31], raising the possibility that the effects we observed were not solely due to the *SOD1*A4V mutation. To address these issues, we used a two-step, zinc-finger nuclease (ZFN)-mediated gene targeting strategy to correct the *SOD1*A4V allele in iPSC line 39b (Figure 2.6A, methods).

Gene-targeting and correction of the *SOD1*A4V mutation was confirmed by sequencing of the genomic locus (Figure 2.6B). To confirm that the genetic correction had eliminated expression of mutant transcripts, we queried the *SOD1* cDNA from both cell lines for the presence of a *Psh*A1 Restriction Fragment Length Polymorphism (RFLP) that results from the missense mutation encoding the SOD1A4V variant. While this PCR-RFLP was clearly detected in cDNA from the 39b-*SOD1*$^{+/A4V}$ cell line, it was eliminated by gene-correction (Figure 2.6C). Quantitative PCR demonstrated that the corrected 39b-*SOD1*$^{+/+}$cell-line did not harbor multiple

insertions of the targeting vector (Figure 2.6D). We also found that *SOD1* transcript levels were unchanged between the parental 39b-*SOD1*$^{+/A4V}$ cell line and its corrected derivative (Figure 2.6E).

Using immunoblotting we examined the SOD1 protein levels in cell extracts from the parental 39b-*SOD1*$^{+/A4V}$ cell line, the intermediate heterozygous knockout line, which expresses only one allele of the *SOD1* gene, and the corrected derivative 39b-*SOD1*$^{+/+}$. We found that while the intermediate 39b-*SOD1*$^{+/-}$ line exhibited approximately half of the SOD1 levels of the parental line, correcting the mutation restored SOD1 protein levels (Figure 2.6F). As expected, based on the lower protein stability of the mutant SOD1A4V variant[32], correction of the *SOD1* mutant allele resulted in a modest increase in SOD1 protein levels relative to those found in parental mutant cells.

Figure 2.6

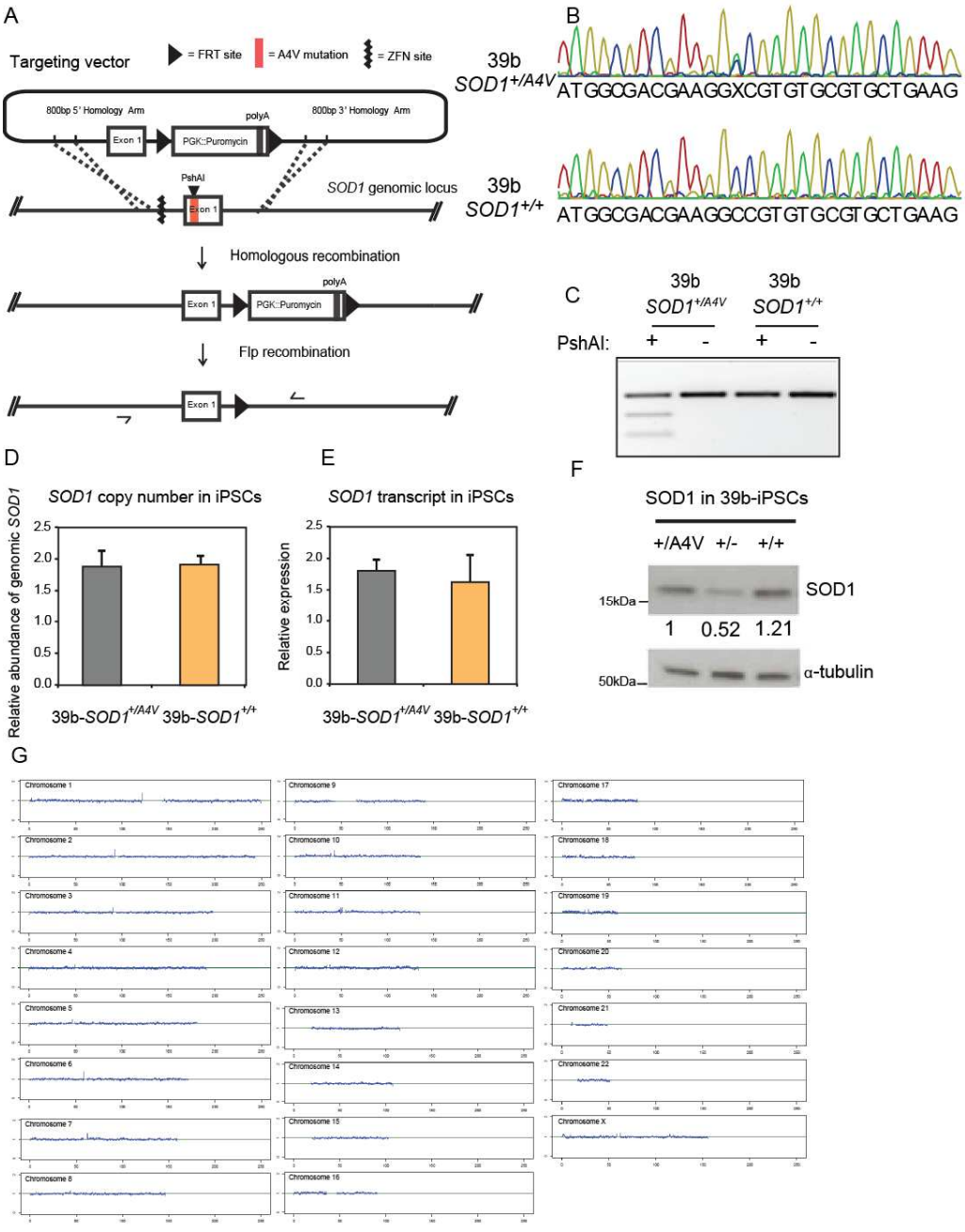
A
Targeting vector
= FRT site
= A4V mutation
= ZFN site
800bp 5' Homology Arm
Exon 1
polyA
PGK::Puromycin
800bp 3' Homology Arm
PshAI
SOD1 genomic locus
Homologous recombination
Flp recombination
B
39b SOD1+/A4V
ATGGCGACGAAGGXCGTGTGCGTGCTGAAG
39b SOD1+/+
ATGGCGACGAAGGCCGTGTGCGTGCTGAAG
C
39b SOD1+/A4V
39b SOD1+/+
PshAI: + - + -
D
SOD1 copy number in iPSCs
Relative abundance of genomic SOD1
39b-SOD1+/A4V
39b-SOD1+/+
E
SOD1 transcript in iPSCs
Relative expression
F
SOD1 in 39b-iPSCs
+/A4V +/- +/+
15kDa
SOD1
1 0.52 1.21
50kDa
α-tubulin
G
Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5
Chromosome 6
Chromosome 7
Chromosome 8
Chromosome 9
Chromosome 10
Chromosome 11
Chromosome 12
Chromosome 13
Chromosome 14
Chromosome 15
Chromosome 16
Chromosome 17
Chromosome 18
Chromosome 19
Chromosome 20
Chromosome 21
Chromosome 22
Chromosome X

Figure 2.6. (Continued) Genetic Correction of the SOD1A4V Mutant Allele (A) Gene targeting strategy used to generate the isogenic 39b-SOD1+/+ iPSC line. Nucleases targeting the SOD1 locus created a double-strand break upstream of Exon1; homologous recombination of the genomic locus with a targeting plasmid with control sequence of Exon1 coupled to PGK::Puromycin replaced the SOD1A4V mutant allele; after antibiotic selection, the resistance cassette was removed by Flp-mediated recombination, leaving only an FRT footprint. ZFN, zincfinger nuclease; FRT, Flippase Recognition Target. (B) Sequencing chromatograms of Exon1 of SOD1 in iPSC line 39b before and after targeting demonstrate the correction of the A4V mutation. (C) PshAI restriction digestion of amplified SOD1 cDNA before and after gene targeting. The mutation creates a PshAI restriction site that is absent in the corrected line. (D and E) qPCR of genomic SOD1 shows that there are no extra copies of the gene in the corrected line (D) and qRT-PCR of cDNA shows that the expression levels of SOD1 are the same in the corrected line (n = 3, ±SEM) (E). (F) SOD1 protein levels assessed by western blot (WB) analysis in parental (39b-SOD1+/A4V), targeted hemizygous knockout (39b-SOD1+/KO), and corrected (39b-SOD1+/+) iPSC clones. (G) To evaluate whether large-scale copy-number alterations had arisen during gene-targeting of the SOD1 locus, we measured sequencing depth of coverage across the genome in 10kb bins for the parental 39b and the corrected 39b control line. Sequencing read depth for the two samples closely matched each other throughout the genome.

Given that undesired genetic variations could potentially arise as result of ZFN off-target activity or clonal expansion, we proceeded to carefully evaluate the genomic sequence and genomic integrity of the parental and genome-edited cell lines. We carried out whole genome sequencing (to 33.4X and 31.1X coverage, for the 39b and gene corrected cell lines, respectively) and analyzed the resulting data with tools including GATK[33] and Genome STRiP[34]. To evaluate whether large-scale copy-number alterations had arisen, we measured sequencing depth of coverage across the genome in 10kb bins, as sequencing read depth (when well normalized) can serve as a proxy for copy number of the underlying genomic DNA[34]. We found that sequencing read depth for the two samples closely matched each other throughout the genome (Figure 2.6G), excluding the possibility that deletions, duplication, or aneuploidy had arisen during passaging, genome editing, or clonal expansion. Furthermore, the top 12,000 genomic loci with sequence similarity to the binding site of the ZFN pair described in this study did not deviate between the parental or gene edited cell lines, indicating highly specific nuclease activity.

To detect other potential genomic sequence changes in the corrected cell line, we compared the fine-scale (SNP and indel) sequence calls between the parental and genome-edited cell lines across their genomes. Overall, we found the corrected cell line to be surprisingly free of such events. However, these analyses were sufficiently sensitive to identify a

likely mitotic recombination event on the q arm of chr12 (from 108Mb to the end of the chromosome). Deeper analysis of this region demonstrated that the event had not induced novel or rare protein-coding variants. Neither had it induced coding variants associated with any known disease state, suggesting this event was likely to be phenotypically neutral.

To evaluate whether either the parental cell line or the genome-edited cell line carried variants associated with ALS, we analyzed the exonic protein-coding sequences of 26 candidate genes implicated in ALS. This analysis identified a single protein-altering sequence difference between the two cell lines, which corresponded to the genome edit of the A4V variant in *SOD1*.

Following the correction of the *SOD1*A4V mutation, we used 39b-*SOD1*$^{+/A4V}$ and 39b-*SOD1*$^{+/A4V}$ cell lines to ask whether the SOD1A4V-encoding variant was necessary for the MN survival and soma size phenotypes observed in the patient-control comparisons. Importantly, correction of the *SOD1*A4V mutation resulted in a significant rescue of both MN survival (n=6, m>13000, P<0.05) and relative soma size deficits (n=3, m=140, P<0.01) (Figure 2.7A-B).

Figure 2.7

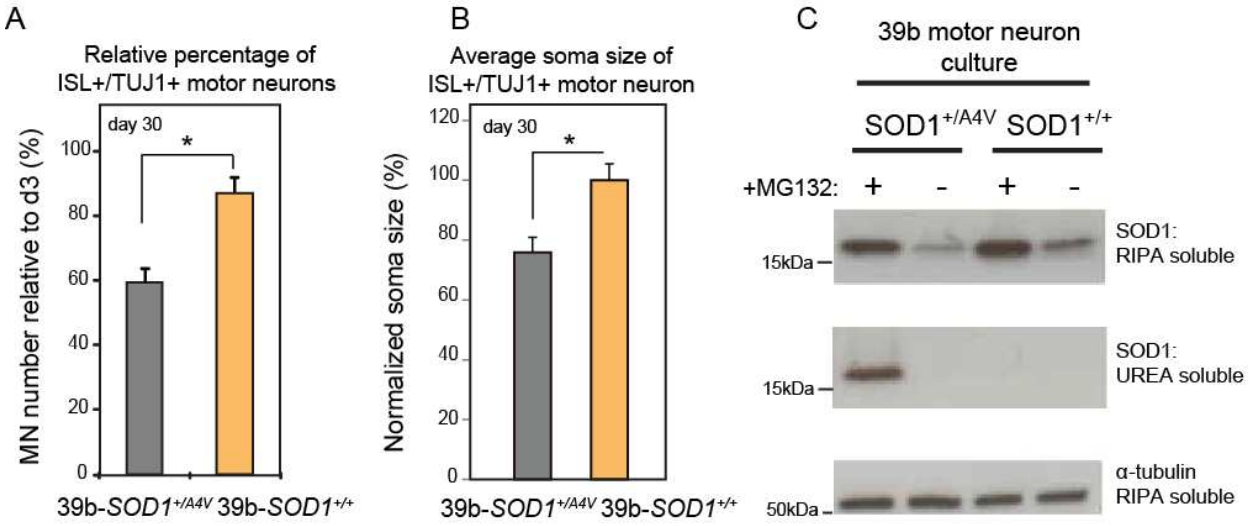

Figure 2.7. (Continued) Genetic Correction of the SOD1A4V Mutant Allele Rescues MN Survival and Soma Size Deficits. (A and B) Isogenic 39b-SOD1+/+ MNs exhibit increased cell survival (n = 6, m > 13,000, ±SEM, $p < 0.05$) (A) and soma size (n = 3, m = 280, ±SEM, $p < 0.01$) (B). (C) WB analysis of SOD1 protein in detergentsoluble (RIPA) and detergent-insoluble (UREA) fractions in day 12 neuronal cultures. Insoluble SOD1 is not detected under normal conditions. After proteasome inhibition by MG132 treatment, insoluble SOD1 selectively accumulates only in SOD1+/A4V MNs and not in the corrected line.

Solubility of Mutant SOD1 in Motor Neuron Cultures. SOD1 protein variants that cause ALS are prone to unfolding, misfolding and ultimately aggregation[35,36]. Aggregation of SOD1 protein and the formation of SOD1-rich inclusions is a pathological finding in animals models of this disease[36], as well as at autopsy in the spinal cords of *SOD1*-linked cases[37] and some sporadic cases[38]. SOD1 aggregation can be detected in spinal cord extracts using detergent solubility assays[39]. However, the relevance of SOD1 insoluble aggregates to disease processes is poorly understood and it remains unresolved whether SOD1 aggregation lies on the critical pathway to neuronal degeneration, and whether these structures are either toxic or protective in nature[38].

The pair of isogenic iPSC lines we developed allowed us to address the state of SOD1 protein in differentiated human spinal MNs expressing physiological levels of wild-type and mutant SOD1. Using immunoblotting assays on detergent-soluble (RIPA) and detergent-insoluble (UREA) fractions obtained from cultures of differentiated $SOD1^{+/+}$ and $SOD1^{+/A4V}$ MNs, we found no evidence of insoluble SOD1 protein under basal culture conditions (Figure 2.7C). In order to sensitize our culture system and validate the ability of our assay to detect insoluble SOD1 protein, we treated MN cultures with the proteasome inhibitor MG132. Treatment with MG132, which inhibits the proteasome by reducing the degradation of ubiquitin-conjugated proteins, increased soluble SOD1 protein levels 2-4 fold. Following MG132 treatment, SOD1 in $SOD1^{+/A4V}$ neurons became detergent-insoluble and was found in the UREA-extracted fraction (Figure 2.7C). Correcting the *SOD1*A4V mutation resulted in elimination of insoluble SOD1 protein in MN cultures following treatment with MG132.

These studies suggest that while it is possible to induce accumulation of insoluble

mutant SOD1 protein in human MNs, insoluble protein did not accumulate to detectable levels under normal culture conditions. We cannot rule out the possibility that undetectable levels of aggregated SOD1 protein were the cause of the increased apoptosis and morphometric changes we found in *SOD1*$^{+/A4V}$ MNs (Figure 2.3-2.4). However, our findings do seem consistent with recent claims that soluble mutant SOD1 can have substantial phenotypic effects[35].

RNA Sequencing of Purified *SOD1*$^{+/A4V}$ and Isogenic Control Motor Neurons. Although the changes in gene expression induced in the mouse nervous system by mutant SOD1 have been extensively studied, it is unknown how mutant SOD1 impacts transcript levels in human MNs. In order to gain insight into the molecular pathways affected by the *SOD1*A4V mutation, we performed RNA sequencing. As in earlier experiments (Figure 2.4-2.7), we differentiated 39b-*SOD1*$^{+/A4V}$ and isogenic control stem cells then plated the resulting MNs on a glial monolayer (Figure 2.8A). We next transduced these cultures with an *Hb9::*RFP lentiviral reporter and used FACS to purify MNs after 15 days of additional culture. RNA and sequencing libraries were prepared from these MNs then subjected to sequencing. We chose day 15 for transcriptional studies as at this time point MNs were physiologically active, but there was only a trend towards reduced MN survival in *SOD1*$^{+/A4V}$ cultures. We reasoned that at this time point we might identify the transcriptional changes that predisposed mature MNs to apoptosis.

The quality of all libraries was validated using FastQC and reads were quantified after alignment and mapping to the reference genome hg19. We first asked whether the effects of the *SOD1*$^{+/A4V}$ genotype on transcriptional patterns in spinal MNs were greater than the variability present between biological replicates of our experimental system. Indeed, unsupervised hierarchical clustering segregated samples based on their *SOD1* genotype, suggesting that effects of the mutant allele were driving measurable transcriptional differences between *SOD1*$^{+/A4V}$ and *SOD1*$^{+/+}$ MNs (Figure 2.8B). As an initial measure of the transcripts most affected by the *SOD1*A4V mutation we determined the identity of the 30 transcripts most

increased and decreased in abundance in mutant MNs relative to controls at a false discovery rate (FDR) of 5% (Figure 2.8C). Analysis of a representative subset of these transcripts by quantitative RT-PCR in independent experiments validated our findings (Figure 2.8D). To understand whether a subset of the gene expression changes found by RNA sequencing were likely to be specific to MNs and not found in other cell types less affected in ALS we determined the expression of a subset of differentially expressed transcripts in the 39b and 39b corrected iPSCs (Figure 2.8E). Of these genes only 19% (5/22) were found to be differentially expressed in both iPSCs and MNs. Furthermore, we performed RNA sequencing on fibroblast cultures isolated from 5 healthy control individuals and the two ALS patients harboring the *SOD1*A4V mutation. While unsupervised hierarchical clustering segregated transcriptomes of MNs harboring the *SOD1*A4V variant as distinct from MNs derived from healthy controls, this was not true for fibroblast samples (Figure 2.8B).

Figure 2.8

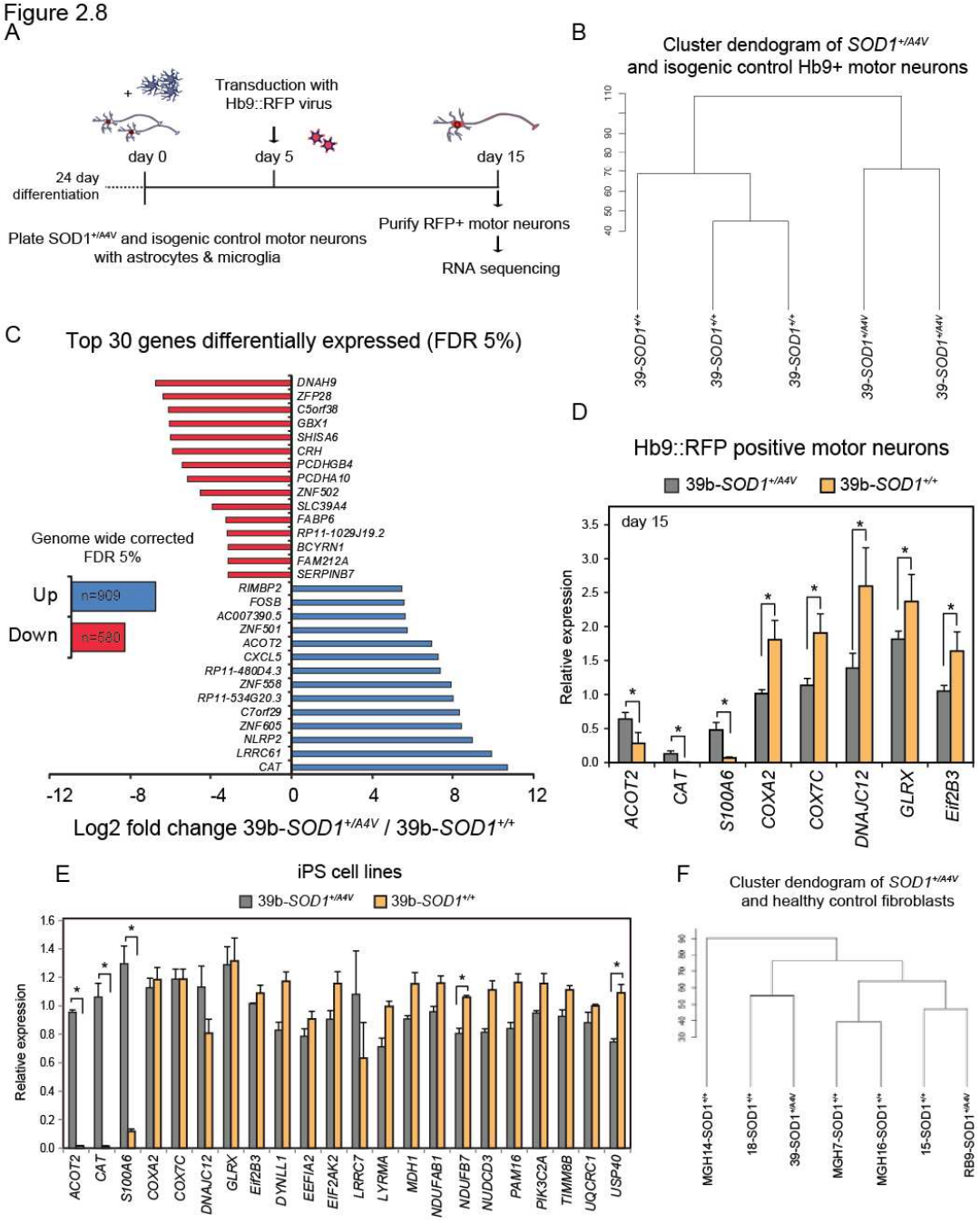
A
Transduction with Hb9::RFP virus
day 0
day 5
day 15
24 day differentiation
Plate SOD1+/A4V and isogenic control motor neurons with astrocytes & microglia
Purify RFP+ motor neurons
RNA sequencing
B
Cluster dendogram of SOD1+/A4V and isogenic control Hb9+ motor neurons
39-SOD1+/+
39-SOD1+/+
39-SOD1+/+
39-SOD1+/A4V
39-SOD1+/A4V
C
Top 30 genes differentially expressed (FDR 5%)
Genome wide corrected FDR 5%
Up
n=909
Down
n=520
DNAH9
ZFP28
C5orf38
GBX1
SHISA6
CRH
PCDHGB4
PCDHA10
ZNF502
SLC39A4
FABP6
RP11-1029J19.2
BCYRN1
FAM212A
SERPINB7
RIMBP2
FOSB
AC007390.5
ZNF501
ACOT2
CXCL5
RP11-480D4.3
ZNF558
RP11-534G20.3
C7orf29
ZNF605
NLRP2
LRRC61
CAT
-12
-8
-4
0
4
8
12
Log2 fold change 39b-SOD1+/A4V / 39b-SOD1+/+
D
Hb9::RFP positive motor neurons
39b-SOD1+/A4V
39b-SOD1+/+
day 15
Relative expression
ACOT2
CAT
S100A6
COXA2
COX7C
DNAJC12
GLRX
Eif2B3
E
iPS cell lines
39b-SOD1+/A4V
39b-SOD1+/+
Relative expression
ACOT2
CAT
S100A6
COXA2
COX7C
DNAJC12
GLRX
Eif2B3
DYNLL1
EEFIA2
EIF2AK2
LRRC7
LYRMA
MDH1
NDUFAB1
NDUFB7
NUDCD3
PAM16
PIK3C2A
TIMM8B
UQCRC1
USP40
F
Cluster dendogram of SOD1+/A4V and healthy control fibroblasts
MGH14-SOD1+/+
18-SOD1+/+
39-SOD1+/A4V
MGH7-SOD1+/+
MGH16-SOD1+/+
15-SOD1+/+
RB9-SOD1+/A4V

Figure 2.8. (Continued) Transcriptional Analysis of 39b-SOD1+/A4V and Isogenic Control MNs by RNA-Seq Reveals Genotypic Regulatory Changes (A) Experimental outline: SOD1+/A4V and isogenic control MN cultures were cocultured with primary glial cells, transduced with an Hb9::RFP lentivirus, and FACS-purified on day 15 for RNA isolation and sequencing. (B) Dendogram demonstrating clear genotypic clustering based on transcriptional changes measured by RNA-seq. (C) Total number of genes and top 30 genes (based on fold change) misregulated in SOD1+/A4V MNs with an FDR of 5%. (D) qRT-PCR validation of misregulated genes in independent samples (n = 3, ±SEM). (E) qRT-PCR analysis in 39b and 39b corrected iPSCs. Genes identified as being differentially regulated as a result of the SOD1 mutation in Hb9::RFP MNs were evaluated in the parental iPSCs. 19% of the genes were found as being significantly different in iPSCs. (F) Cluster dendogram of RNA sequencing from SOD1A4V and healthy sex-matched fibroblast samples. The A4V fibroblasts do not segregate from the healthy controls.

Validating the importance of this unbiased approach, the vast majority of genes we identified to be differentially expressed between *SOD1*$^{+/A4V}$ and control MNs had not previously been implicated as being modulated by mutant SOD1. These included *ACOT2*, an enzyme that hydrolyzes Coenzyme A (CoA) esters[40]; the transcription factor *FOSB* which regulates BDNF expression[41], and *NLRP2,* a NOD-like receptor, pro-apoptotic component of the inflammasome[42]. Amongst the most-downregulated genes in *SOD1*$^{+/A4V}$ MNs, we identified a transcript that encodes the molecular motor protein DNAH9. Transcripts of *CXCL5*, a chemokine indicative of an inflammatory response previously shown to be stimulated by mutant SOD1 protein[42] were similarly increased here. We also found that corticotropin-releasing hormone (*CRH*), which has been implicated in other neurodegenerative diseases[43] and found to be expressed at lower levels in ALS patients was significantly reduced in *SOD1*$^{+/A4V}$ MNs.

Ontology of transcripts modulated in *SOD1*$^{+/A4V}$ Motor Neurons. In order to probe the RNA sequencing data for further biological meaning we utilized two bioinformatics tools that query for enriched gene ontology terms. We first performed gene-annotation enrichment analysis with DAVID[44], using all the genes that were significantly altered (909 upregulated and 580 downregulated) in *SOD1*$^{+/A4V}$ MNs at a FDR of 5%. A total of 27 and 65 gene terms were enriched when increased and reduced transcripts were considered respectively (FDR<5%). Transcripts implicated in 'cytoskeleton' organization (rank of significance 5, FDR=5.84×10^{-7}), where amongst the most significantly induced in *SOD1*$^{+/A4V}$ MNs, consistent with the morphological alterations that we observed in these cells relative to isogenic controls (Figure

2.7B). Transcripts involved in 'transcriptional regulation' (rank of significance 13, FDR=$1.7x10^{-4}$) and 'motor proteins' (rank of significance 18, FDR=$6.4x10^{-3}$), were also induced as a result of the *SOD1* mutation.

Gene ontology analysis revealed that amongst the significantly decreased transcripts in *SOD1*$^{+/A4V}$ MNs, there was a very strong enrichment for genes implicated in mitochondrial function and structure. In particular, 60% of all downregulated ontology terms were related to mitochondria. A total of 93 genes curated in the term 'mitochondrion' (rank of significance 1, FDR=$6.3x10^{-26}$). In addition, genes implicated in 'translation' (rank of significance 34, FDR=$1.3x10^{-4}$) were also repressed.

As an alternative approach for querying our RNA sequencing data, we performed Gene Set Enrichment Analysis (GSEA)[45,46]. GSEA identified 16 gene sets that were significantly induced in *SOD1*$^{+/A4V}$ MNs (NES<1.5). Amongst these gene sets, were the motor proteins *'kinesins'* (rank of significance 5; NES=1.7). GSEA also identified 100 gene sets to be significantly repressed in *SOD1*$^{+/A4V}$ MNs (NES<1.5). Notably, gene sets associated with mitochondrial function and translation were again amongst the most significantly suppressed. In particular, gene sets annotated as being involved in 'peptide chain elongation' (rank of significance=2; NES=-3.45), 'SRP-dependent co-translational protein targeting to membrane' (rank of significance=3; NES =-3.45), '3-UTR mediated translational regulation' (rank of significance 5; NES =-3.42), and 'translation' (rank of significance=8; NES=-3.31) were all significantly suppressed.

***SOD1*$^{+/A4V}$ Motor Neurons Exhibit Disturbances in Mitochondrial Morphology and Motility.** To determine whether the transcriptional changes in mitochondrial genes that we identified by RNA-Seq in *SOD1*$^{+/A4V}$ MNs were indicative of actual disturbances to mitochondria, we performed electron microscopy (EM) studies (Figure 2.9B). Whereas mitochondrial morphology was normal in MNs derived from a control cell line (18a), mitochondria in *SOD1*$^{+/A4V}$ MNs (39b and Rb9d) were commonly deranged and more vacuolar in appearance. These

differences were mostly apparent in neuronal processes. We concluded that distortion in mitochondrial morphology was mediated by expression of the *SOD1*A4V mutant allele as correction of the mutation eliminated this phenotype (Figure 2.9B). To further validate mitochondrial damage in *SOD1*$^{+/A4V}$ MNs, we used immunoblotting assays to quantify the levels of two mitochondrial proteins, SDHA (SUCCINATE DEHYDROGENASE COMPLEX SUBUNIT A), which is encoded in the nucleus, and MT-COX1 (CYTOCHROME C OXIDASE I), encoded by mitochondrial DNA. Consistent with mitochondrial defects mediated by expression of mutant SOD1, we found that correction of the *SOD1*A4V allele in 39b-*SOD1*$^{+/+}$ MNs increased the protein levels of both SDHA and MT-COX1, relative to the parental 39b-*SOD1*$^{+/A4V}$ MNs.

Figure 2.9

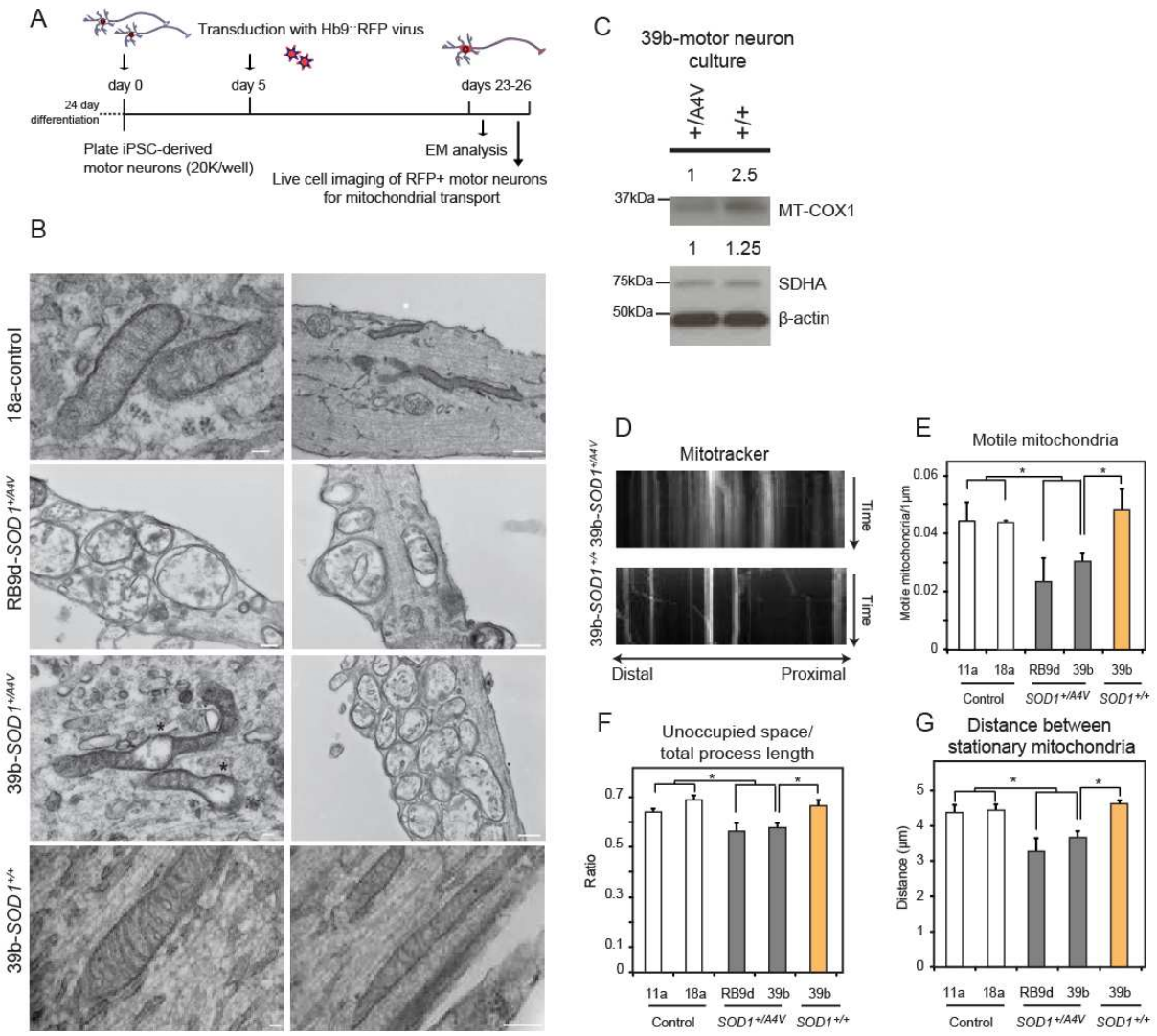

Figure 2.9. (Continued) SOD1+/A4V MNs Exhibit Mitochondrial Defects that Are Rescued by Genetic Correction of the SOD1A4V Mutation (A) Experimental outline: SOD1+/A4V and isogenic control MN cultures were plated at low densities, transduced with an Hb9::RFP lentivirus on day 5, and analyzed on days 23–26. (B) EM analysis demonstrating that in contrast to control neurons (top panel), SOD1+/A4V neurons (middle panels) often exhibited swelling, morphological disorganization, and clustering in neurites. Mitochondria in genetically corrected neurons (bottom panel) exhibited normal morphological characteristics. Representative images out of three independent experiments are shown. Scale bar, 100 nm. (C) WB analysis of mitochondrial proteins. (D–G) Representative kymographs (D) of MitoTracker-stained mitochondria in MNs after live cell imaging. Distal and proximal orientation relative to cell body is indicated. SOD1+/A4V MNs exhibit mitochondria transportation deficiencies relative to controls with (E) fewer motile mitochondria, (F) less space unoccupied by mitochondria relative to axon length, and (G) shorter distances between stationary mitochondria (n = 4, ±SEM, $p < 0.05$).

As changes in mitochondrial motility have been reported previously in mouse models of

ALS we next sought to observe the movements of mitochondria within the axons of our *in vitro*-derived MNs (Figure 2.9D-G). MN cultures differentiated from control and *SOD1*$^{+/A4V}$ iPSCs were co-labeled with *Hb9::*RFP and MitoTracker-Green, which selectively stained mitochondria regardless of their membrane potential. We then carried out live cell imaging to register the movement of mitochondria along MN processes over the course of 5 minutes, and generated time/distance kymographs for further analysis (Figure 2.9D). We found that the *SOD1*A4V mutation resulted in a significant decrease in the number of motile mitochondria (Figure 2.9E). This decreased motility was coupled to an increase in mitochondrial density in processes, as measured by the shorter distance between stationary mitochondria and the significantly smaller amount of space unoccupied by mitochondria (Figure 2.9F,G).

***SOD1*$^{+/A4V}$ Motor Neurons Exhibit Signatures of an Unfolded Protein Response and ER Stress Induction**. To gain insight into how changes in individual transcripts annotated in the 'translation' gene set contributed to the GSEA analysis, we plotted their levels in *SOD1*$^{+/A4V}$ relative to isogenic control MNs (Figure 2.10A). Strikingly, we found that 93% of transcripts curated into this gene set were reduced in *SOD1*$^{+/A4V}$ MNs. Amongst these there were several subunits of eukaryotic translation initiation factors (*EIF2B3*, *EIF3K*). Translational inhibition is well-established as a hallmark of ER stress and the unfolded protein responses (UPR). We therefore investigated whether this pathway was activated[47].

The UPR is a signaling cascade that is activated in response to accumulation of misfolded proteins and while it is thought to be initially cytoprotective, its persistent activation can in turn lead to apoptosis[48,49]. The UPR is initiated though its proximal sensors, which include the proteins PERK and IRE1 (Figure 2.10B). In one of the branches of the pathway, PERK directly leads to an increase in the phosphorylation of eukaryotic translation initiation factor 2 subunit a (EIF2a).

In order to decrease the protein load of the cell, pEIF2a leads to a global attenuation of translation and to selective translation of the transcription factor ATF4[47]. In another branch of

the UPR, IRE1 cleaves the mRNA of *XBP1,* creating an active form of the transcription factor (sXBP1)[49]. Spliced XBP1 along with ATF4 modulate expression of multiple downstream effectors of the UPR, including chaperone proteins. Importantly, RNA-seq analysis identified the heat-shock proteins *DNAJC12* and *HSBP1,* the prefoldin subunits *PFDN2* and *PFDN5,* as well as the chaperonin subunits *CCT4* and *CCT7* as being differentially expressed in *SOD1*$^{+/A4V}$ MNs. Consistent with the transcriptional changes reflecting an UPR in our patient-specific *SOD1*$^{+/A4V}$ MNs, ER-stress pathways and the UPR are activated in rodent ALS models and patient spinal cords[50-52].

To more directly test whether the transcriptional signature we observed was caused by activation of the UPR in our MN cultures, we examined the levels of phopho-EIF2a. Using immunoblotting we found increased levels of pEIF2a in *SOD1*$^{+/A4V}$ MNs on days 7, 9 and 12 (on average 1.5-fold higher than isogenic controls) (Figure 2.10C). We also found that the spliced *XBP1* transcript was significantly elevated in *SOD1*$^{+/A4V}$ MNs relative to isogenic controls (n=3, $P<0.05$) (Figure 2.10D), suggesting that this arm of the UPR pathway was also activated. In order to test if ER stress directly contributed to mutant SOD1-mediated toxicity in our cell culture system, we genetically manipulated the two UPR branches using siRNA knockdown of *XBP1* and *ATF4* and assessed the effect on the survival of mutant and controls MNs (Figure 2.10E-F). Knockdown of the *XBP1* transcript resulted in a modest but significant increase in the survival of *SOD1*$^{+/A4V}$ MNs (Figure 2.10G). In contrast, knockdown of *XBP1* in the isogenic controls did not lead to an increase in their survival, but instead there was a trend for a reduced MN number, suggesting that *XBP1* might provide a protective function in this context. Knockdown of *ATF4* transcript levels depressed survival of both *SOD1*$^{+/A4V}$ and *SOD1*$^{+/+}$ MNs, implying that this protein plays an important protective role in both contexts (Figure 2.10G). Given that reducing *ATF4* levels was detrimental to the survival of both *SOD1*$^{+/A4V}$ and control MNs we asked whether a further induction of pEIF2a would confer protection. Salubrinal is a selective inhibitor of phosphatases, which dephosphorylate pEIF2a[53]. Furthermore Salubrinal extends the survival

of the *SOD1*^{G93A} mouse model by reducing the UPR and increasing innervation of neuromuscular junction (NMJs)[52]. Treatment of MN cultures with salubrinal from d15-30 led to a modest but significant increase in the survival of *SOD1*$^{+/A4V}$ MNs, whereas the survival of controls was not affected (n=2, m>10000, P<0.05) (Figure 2.10H).

Taken together these data clearly demonstrate that the UPR is activated in human MNs by the A4V mutation. Furthermore, while ATF4 seems to be generally protective in both contexts, treatment with a compound that inhibits EIF2a phosphatases improves survival of *SOD1*$^{+/A4V}$ MNs. In contrast, modulation of the *XBP1* branch of the UPR pathway had a more complicated effect. Knockdown of *XBP1* improved *SOD1*$^{+/A4V}$ MN survival suggesting that chronic activation of this arm of the UPR may play a role in the death of the mutant MNs.

Figure 2.10

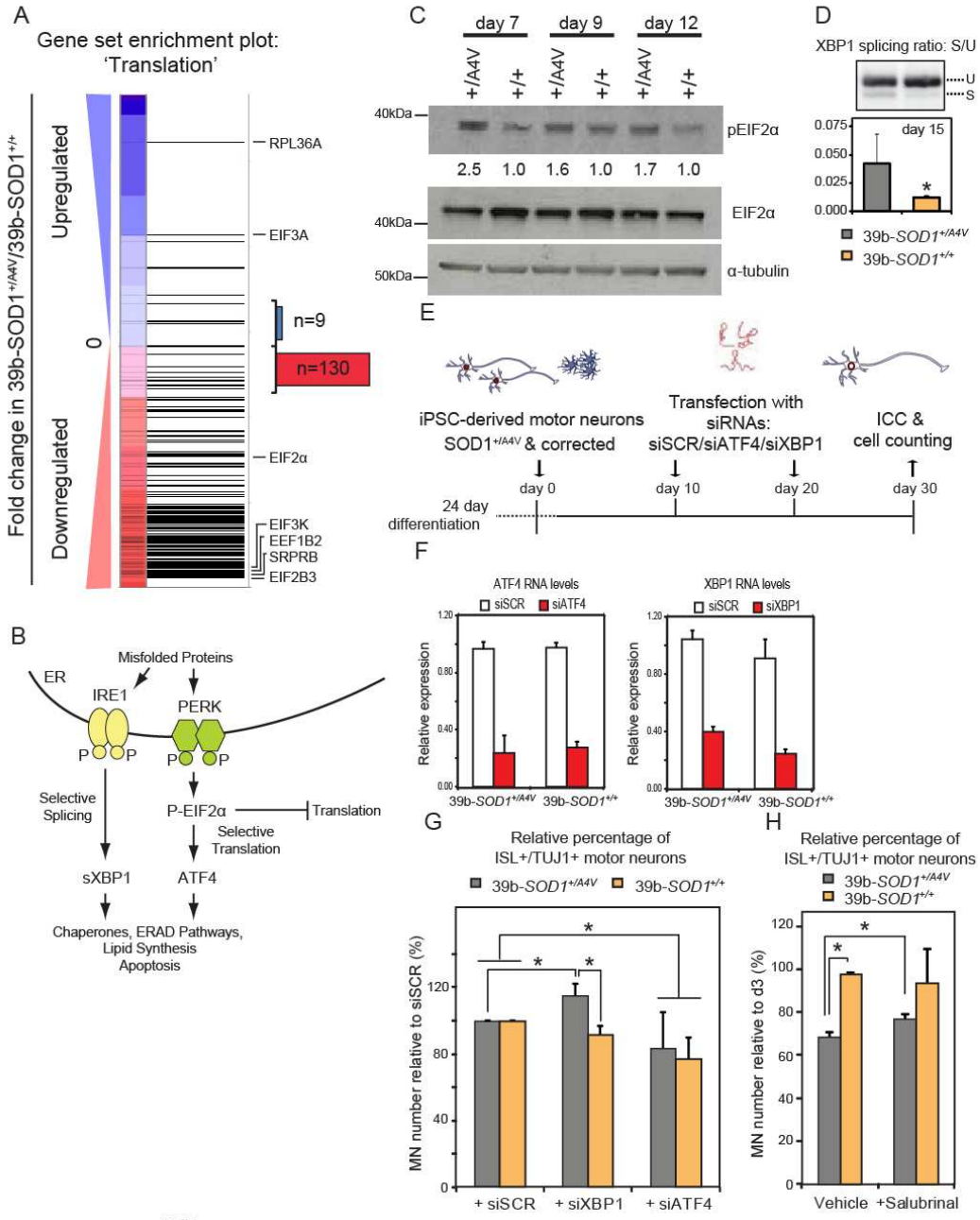
A
Gene set enrichment plot:
'Translation'
Fold change in 39b-SOD1+/A4V/39b-SOD1+/+
Upregulated
0
Downregulated
RPL36A
EIF3A
EIF2α
EIF3K
EEF1B2
SRPRB
EIF2B3
n=9
n=130
B
ER
Misfolded Proteins
IRE1
PERK
P
Selective Splicing
P-EIF2α
Translation
Selective Translation
sXBP1
ATF4
Chaperones, ERAD Pathways, Lipid Synthesis Apoptosis
C
day 7
day 9
day 12
+/A4V
+/+
40kDa
pEIF2α
2.5 1.0 1.6 1.0 1.7 1.0
EIF2α
50kDa
α-tubulin
D
XBP1 splicing ratio: S/U
U
S
day 15
0.075
0.050
0.025
0.000
*
39b-SOD1+/A4V
39b-SOD1+/+
E
iPSC-derived motor neurons SOD1+/A4V & corrected
Transfection with siRNAs: siSCR/siATF4/siXBP1
ICC & cell counting
day 0
day 10
day 20
day 30
24 day differentiation
F
ATF4 RNA levels
siSCR
siATF4
Relative expression
39b-SOD1+/A4V
39b-SOD1+/+
XBP1 RNA levels
siSCR
siXBP1
Relative expression
39b-SOD1+/A4V
39b-SOD1+/+
G
Relative percentage of ISL+/TUJ1+ motor neurons
39b-SOD1+/A4V
39b-SOD1+/+
MN number relative to siSCR (%)
120
80
40
0
*
*
*
+ siSCR
+ siXBP1
+ siATF4
H
Relative percentage of ISL+/TUJ1+ motor neurons
39b-SOD1+/A4V
39b-SOD1+/+
MN number relative to d3 (%)
120
100
80
60
40
20
0
*
*
Vehicle
+Salubrinal

Figure 2.10. (Continued) SOD1+/A4V MNs Exhibit Signatures of a UPR and Are Selectively Vulnerable to ER Stress Induction (A) GSEA of transcriptional changes in SOD1+/A4V MNs shows strong downregulation (normalized enrichment score: 3.31) for genes involved in translational capacity. Horizontal black bars represent individual genes with representative examples indicated. Blue or red represents genes that were upregulated or downregulated, respectively, in SOD1+/A4V relative to isogenic control MNs. Out of 139 genes detected that are annotated as involved in regulation of translation, 130 were downregulated, consistent with activation of the UPR pathway. (B) Diagram illustrating the canonical UPR. (C) WB analysis demonstrates increased levels of phosphorylated EIF2a, a marker for activation of the UPR pathway, in SOD1+/A4V MN cultures. The percentages relative to control samples for each time point are shown. (D) SOD1+/A4V MNs exhibit increased levels of XBP1 splicing, a marker of ER stress. RNA was isolated from purified Hb9::RFP MNs after 15 days in culture ($n = 3$, ±SEM, $p < 0.05$). U, unspliced; S, spliced. (E) Experimental strategy used to assess the contribution of XBP1 and ATF4 in the survival of SOD1+/A4V MNs. (F) Knockdown of ATF4 and XBP1 in human motor neuron cultures after 30 days. (G) SOD1+/A4V MN numbers selectively increase after XBP1 knockdown, while ATF4 knockdown significantly decreases numbers in both control and SOD1+/A4V cases ($n = 3$, $m > 9,800$ and $m > 8,500$, ±SEM, $p < 0.05$). (H) Salubrinal has a modest but positive effect on survival of SOD1+/A4V MNs ($n = 2$, $m > 10,000$, ±SEM, $p < 0.05$).

ER Stress is Inherent in Human Motor Neurons and Spinal Cord. Although the *SOD1*A4V mutation increased the levels of pEIF2a and of the spliced form of *XBP1*, we noted that, surprisingly control MNs also expressed these markers of an activated UPR, albeit at a lower level. This observation led us to hypothesize a novel and relatively simple explanation for the selective sensitivity of MNs to the effects of the A4V mutation, which is that MNs in general are more sensitive to unfolded proteins and ER stress because the presence of these liabilities is a normal and perhaps even constitutive aspect of this neuronal sub-type's inherent biology.

To test this hypothesis, we compared the levels of *XBP1* splicing in purified, control HUES3 *Hb9*::GFP+ MNs with levels found in a range of cell types including astrocytes, ES cells, neural progenitors, fibroblasts, *Hb9::*GFP-negative neurons and ES-derived anterior neurons. Of all the cell types examined, only the control MNs displayed detectable levels of spliced *XBP1* (Figure 2.11A). *XBP1* splicing was not found at early time points but as the MNs matured in culture, it became evident. To test the *in vivo* relevance of these findings, we compared the levels of *XBP1* splicing in RNA isolated from control human brain and control spinal cord (n=2; each replicate was pooled RNA from 22, non-overlapping, healthy controls). Strikingly, we saw little or no evidence for *XBP1* splicing in the brain, while in the control spinal cord there was a marked and significant accumulation of the spliced transcript (Figure 2.11A).

We wondered whether this ongoing activation of the UPR pathway was associated with

an increased sensitivity to agents such as DTT that unfold proteins and increase levels of ER and oxidative stress. To test this, we prepared control MNs, fetal astrocytes and fibroblasts and subjected them to an acute challenge with DTT. Strikingly, MNs exhibited a greater susceptibility to DTT administration than the other cell types (Figure 2.11B) ($n=3$, $P<0.05$). Given that the largest A4V MNs were most predisposed to cell death (Figure 2.3K), we wondered whether these MNs would also be more sensitive to the effects of DTT. When we analyzed the area of the soma of MNs we found that the average soma size decreased substantially after treatment (Figure 2.11C), which given the acute nature of the treatment seemed most consistent with death of the largest MNs.

These observations prompted us to investigate whether MN size correlated with basal ER stress levels. To address this question we generated control *Hb9::*RFP MNs and after 15 days, FACS-purified them as well as separated large and small MN populations (Figure 2.11D). The effectiveness of this approach in separating MNs based on size was validated by re-plating a sub-set of the FACS-purified MNs and measuring their soma sizes (Figure 2.11E). We then isolated RNA from the large and small populations and examined the levels of spliced *XBP1* as an indicator of ER stress. Larger MNs showed significantly higher levels of spliced *XBP1* than smaller ones (Figure 2.11F) suggesting that an increased constitutive ER stress may contribute to their increased vulnerability in ALS.

Figure 2.11

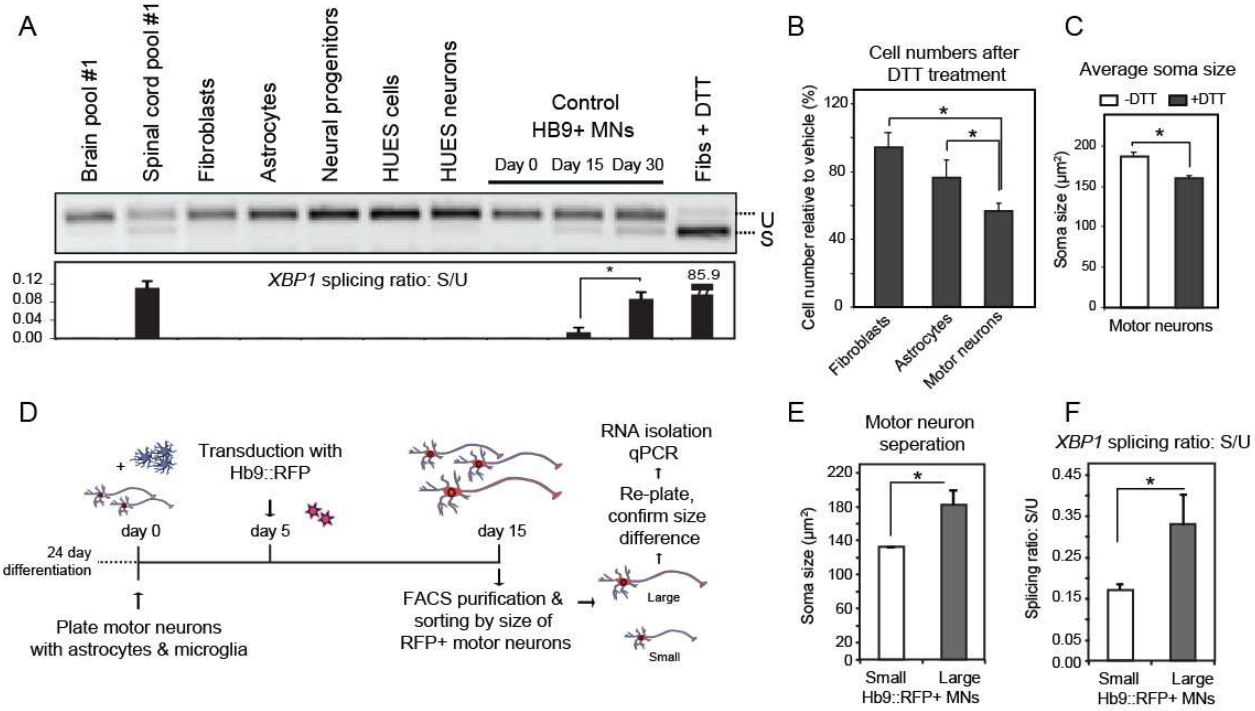

Figure 2.11. (Continued) Human MNs Exhibit Increased Levels of Basal ER Stress that Are Dependent on Their Physiological Activity. (A) Purified Hb9::GFP control MNs show higher levels of spliced XBP1 relative to other human cell types. Human spinal cord RNA also shows higher levels relative to brain RNA. U, unspliced; S, spliced. (B and C) Control MNs are more vulnerable to acute ER stress induction (DTT, 2 mM) relative to fibroblasts or astrocytes (B), while (C) DTT treatment leads to a reduction in MN soma size (n = 2, ±SD, $p < 0.05$). (D) Experimental strategy used to isolate MNs based on cell size. Control MN cultures were infected with the Hb9::RFP virus on day 5. On day 15, RFP+ MNs were FACS purified and sorted by size to achieve separation of small and large cells. A subset of cells (E) were replated to ensure confirmation of soma size by the measurement of MAP2+ cell bodies. (F) Basal levels of XBP1 splicing were assessed in RNA isolated from the remaining purified MNs, which showed that larger MNs had higher levels of spliced XBP1 than smaller ones did (n = 3, ± SEM, $p < 0.05$).

Hyperexcitability in Motor Neurons expressing SOD1 A4V. Previous nerve conduction studies evaluating axonal threshold (strength-duration time constant and recovery cycle times) in ALS patients demonstrated increased axonal membrane excitability[54,55], and the degree of hyperexcitability correlated with ALS patient survival[56]. These data indicate that increased membrane excitability may be important as a contributor to the disease, and modeling suggests that either increased persistent sodium or reduced delayed-rectifier potassium currents could be responsible for the axonal hyperexcitability[55,57]. However, whether the excitability results from changes in the autonomous properties of the motor neurons themselves

cannot be determined by this technique[58].

To test the excitability of our stem cell derived MNs we used multi-electrode arrays (MEAs) to measure the firing properties of our cultures. Because substantial motor neuron death begins in ALS motor neurons after 15 days of neuronal maturation in our culture conditions, we compared MEA recordings of $SOD1^{+/+}$ and $SOD1^{+/A4V}$ motor neurons after a shorter culture time (14 days), to avoid the possibility that increased firing might reflect either neuronal death or a select group of surviving hyperexcitable neurons. While the baseline spike rate was lower at 14 compared to 28 days, $SOD1^{+/A4V}$ neurons had a far higher spontaneous firing rate than $SOD1^{+/+}$ neurons ($p=0.01$ for total rate, t-test; Figure 2.12A; average mean firing rate 1.30 ± 0.10 Hz for n=122 $SOD1^{+/+}$ and 1.50 ± 0.08 Hz for n=208 $SOD1^{+/A4V}$; $p<0.05$, t-test). Because the genetically corrected controls were less excitable, we conclude the hyperexcitability phenotype reflected the necessity of the disease-initiating mutation and preceded the progressive motor neuron death. After establishing the hyperexcitability of motor neurons expressing the SOD1 A4V variant we wanted to evaluate the robustness of this phenotype in other familial forms of ALS. To accomplish this goal we used MEAs to record the firing properties of MNs derived from patients with SOD1, FUS, and C9orf72 mutations as well as a large group of controls. We found the spontaneous action potential firing was significantly increased in all mutant cases (Figure 2.12D ANOVA $p < 10^{-7}$; Tukey's post hoc test for control versus SOD1 $p < 0.01$, control versus C9orf72 $p < 0.01$, control versus FUS $p < 0.05$) demonstrating the broad nature of the increased excitability phenotype in familial ALS.

The MEA recordings demonstrated an important physiological phenotype in MNs derived from ALS patient iPSCs. We therefore wanted to probe the neuronal firing properties with patch clamp techniques to gain a more in depth understanding of the mutant MNs hyperexcitability. When we compared the motor neurons derived from the $SOD1^{+/A4V}$ and their isogenic control cell lines in three parallel patch experiments, we observed a marked difference in the number of action potentials elicited during ramp depolarization ($p<0.05$, Mann-Whitney U test; Figure

2.12B). There was substantial variability in the number of action potentials in motor neurons from the same line tested across multiple differentiations, but the increased number of action potentials in ALS motor neurons relative to control motor neurons was preserved. This result underscores the importance of performing repeated parallel differentiations in which equal numbers of control and ALS motor neurons are analyzed from each differentiation.

Figure 2.12

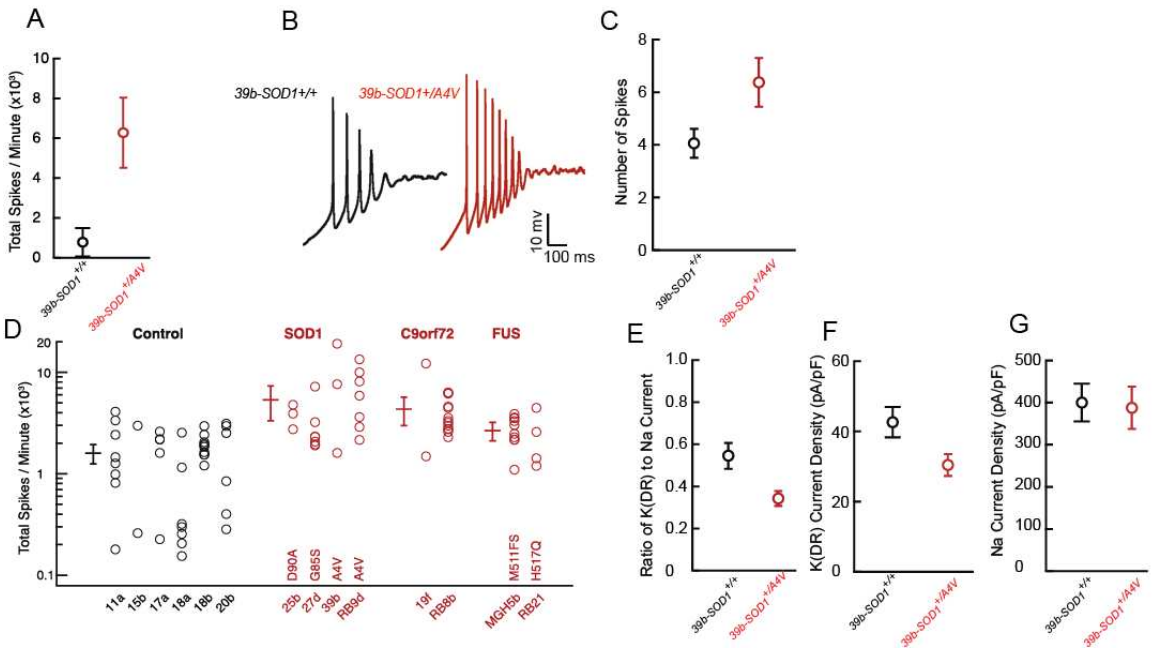

Figure 2.12 (Continued) Physiological recordings reveal increased excitabiilty in SOD1A4V/+ MNs.(A) Total action potential firing rate during 1 min of recording from MEAs cultured for 14 days on the arrays (SOD1+/+, n = 4; mean 775 ± 712 spikes/min; 39b, n = 4; SOD1+/A4V mean 6,278 ± 1,758 spikes/min; p = 0.01, ttest). Error bars represent SEM. (B) Representative current clamp recordings during ramp depolarization from control and ALS patient-derived motor neurons. (C) Separate experiments showing average number of action potentials during ramp depolarization from SOD1+/+ (n=17; mean 4.1±0.5) and SOD1+/A4V (n=19; mean 6.4±0.9) motor neurons from three additional differentiations (p<0.05, Mann-Whitney U test). (D) Multielectrode array recordings of motor neurons derived from control (11a, n = 8; 15b, n = 2; 17a, n = 5; 18a, n = 7; 18b, n = 10; 20b, n = 6), SOD1 (25b, D90A, n = 3; 27d, G85S, n = 7; 39b, A4V, n = 3; RB9d, A4V, n = 7), C9orf72 expansion (19f, n = 2; RB8B, n = 13) and FUS (MGH5b, frame shift mutation at residue 511, n = 10; RB21, H517Q, n = 4) subjects cultured for 14 days. ANOVA, p < 10-7; Tukey's post hoc tests for control versus SOD1 p < 0.01, control versus C9orf72 p < 0.01, control versus FUS p < 0.05. For subject 18, motor neurons from two different iPSC lines were recorded. Error bars are 95% confidence interval. (E) Experiments from three separate differentiations showing averaged delayed-rectifier steady-state potassium current amplitude relative to peak sodium current amplitude in SOD1+/+ (n=18; mean 0.54±0.061) and SOD1+/A4V (n = 19; mean 0.32 ± 0.036; p < 0.005, ttest). (F) Direct measurement of delayed-rectifier voltage-gated potassium current isolated by holding at 30mV, stepping to a test-potential of +40mV for 2s and normalizing steady state current amplitude to cell capacitance in SOD1+/+ (n=19; mean 42.6 ± 4.3 pA/pF) and SOD1+/A4V (n=18; mean 30.3 ± 3.1 pA/pF; p<0.05, ttest) derived motor neurons using cells from two additional separate differentiations. (G) Peak sodium current amplitude normalized to cell capacitance in SOD1+/+ (n=16; mean 400.4 ± 44.7 pA/pF) and SOD1+/A4V (n=15; mean 387.1 ± 50.5 pA/pF; p=0.8, ttest) derived motor neurons. Error bars represent SEM.

In addition to quantifying the electrical excitability of individual neurons, patch clamp recording enables quantitative investigation of the specific currents that determine excitability. To identify the electrophysiological mechanisms responsible for the increased firing of the mutant motor neurons, we performed voltage-clamp experiments using *HB9*::RFP-positive motor neurons to examine different current components. As an index of excitatory and inhibitory voltage-dependent ion channels, we quantified the ratio of outward delayed-rectifier potassium current to inward transient sodium current. In four repeated differentiations of motor neurons from the control and *SOD1$^{+/A4V}$* isogenic iPSC lines, we observed that the ratio of delayed-rectifier potassium to transient sodium current was consistently smaller in the *SOD1$^{+/A4V}$* motor neurons (p<0.001, t-test; Figure 2.12D-F). The difference appeared to be driven primarily by the reduced delayed-rectifier potassium channel component, as the difference in steady-state potassium current amplitude normalized to individual cell capacitance between SOD1$^{+/A4V}$ and SOD1$^{+/+}$ yielded an increase in the relative steady-state delayed-rectifier potassium current amplitude in the SOD1$^{+/+}$ derived MNs (p<0.005, t-test; Figure 2.12D). We found a marked reduction in delayed-rectifier current magnitude in the SOD1$^{A4V/+}$ compared to their isogenic control motor neurons (p< 0.05, t-test; Figure 2.12E) but no difference in measured sodium

current peak amplitudes between the motor neurons (p=0.8, t-test; Figure 2.12F). Since voltage-gated potassium channels repolarize the membrane potential back to negative values after an action potential, a decrease in such currents likely contributes to the observed increased action potential firing in the ALS motor neurons.

Thus, correction of the deficit in delayed-rectifier potassium current in $SOD1^{+/+}$ motor neurons may enable repolarization of the membrane potential back to normal hyperpolarized values and a reduction of the excitability to levels found in wild-type motor neurons.
Like most mammalian neurons, motor neurons express many different types of voltage-activated potassium channels and pharmacological dissection and quantification into distinct components is challenging. Regardless of which currents are altered to produce hyperexcitability in the diseased motor neurons, Kv7 (KCNQ) channels are particularly attractive targets for correcting the hyperexcitability, because of their activation at subthreshold voltages and subsequent powerful control of excitability[59]. Given this consideration and our finding of reduced delayed-rectifier potassium currents in the ALS-derived motor neurons, we hypothesized that retigabine, a specific activator of subthreshold Kv7 currents that is an effective, clinically-approved anticonvulsant[60], might block hyperexcitability in the *SOD1*$^{+/A4V}$ motor neurons. In whole-cell patch clamp, we found that retigabine significantly increased the minimal current step necessary for action potential generation (rheobase) by 3.6 ± 2.4 pA ($p<0.05$, Wilcoxon signed rank test; Figure 2.13A). Retigabine also stopped the spontaneous firing of *HB9*::RFP-positive motor neurons and hyperpolarized the resting membrane potential by 6.0 ± 2.2 mV (p=0.001, t-test; Figure 2.13B). Because these experiments were performed in the presence of blockers of glutamatergic, GABAergic, and glycinergic receptors, the reduction in excitability caused by retigabine likely resulted from a direct effect on motor neurons. We used MEA recordings to determine a dose-response curve for inhibition of spontaneous firing by retigabine of *SOD1*$^{+/A4V}$ ALS-derived neurons. Retigabine suppressed spontaneous firing in ALS neurons with an EC_{50} of 1.5 µM (Figure 2.13C), a concentration consistent with its

pharmacological activity as an anti-epileptic agent in patients and similar to its EC_{50} for Kv7 channels[61]. In line with this finding, analysis of the RNA-Seq data from the FACS-sorted motor neurons confirmed expression of Kv7 channels.

Figure 2.13

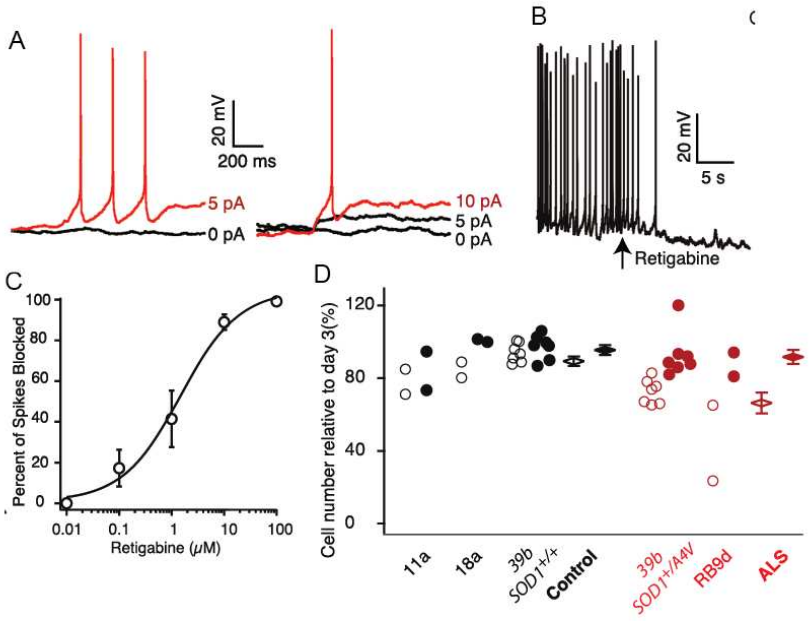

Figure 2.13. Retigabine Reduces Motor Neuron Excitability and Increases Survival. (A) Rheobase measurements in a 39b Hb9::RFP-positive ALS-derived motor neuron in whole-cell patch clamp before (left) and after (right) the application of 10mM retigabine (base line rheobase 4.8 ± 1.5 pA versus post-retigabine rheobase 8.4 ± 2.2 pA; n = 11; $p < 0.05$, Wilcoxon signed rank test). (B) Representative current clamp recording showing effect of 10mM retigabine on membrane voltage and spontaneous firing (baseline Vm 60.4 ± 2.9 mV versus post-retigabine Vm 66.3 ± 3.6 mV, n = 11; p = 0.001, ttest). In (A) and (B), CNQX (15 mM), D-AP5 (20 mM), bicuculline (25 mM), and strychnine (2.5 mM) were added to the external solution. (C) Dose-response curve for retigabine on suppression of spontaneous action potentials in MEA recording and Hill plot fit of mean data from 39b (n = 4) and RB9d (n = 4) with EC50 1.5 ± 0.8 mM. (D) Effect of vehicle (open circles) and 1 mM retigabine (filled circles) treatment from days 14–28 of culture on the survival of Islet-positive, Tuj1-positive motor neurons measured at day 30 (total control n = 11; total ALS n = 9; F-test for effect of retigabine on all cells p = 3.83 X 10-4; effect of retigabine in ALS motor neurons, red, 25.3% (SD 5.6; ttest p = 6.43 X 10-5); effect of retigabine in control motor neurons, black, 6.1% (SD 5.1, p = 0.23). Cell counts are from individual wells for four separate differentiations. Error bars represent SEM.

The efficacy of retigabine allowed us to evaluate the possibility that hyperexcitability is a causal factor or upstream modulator of motor neuron degeneration in ALS. We tested the effect of retigabine on the survival of control and *SOD1*$^{+/A4V}$ motor neurons over the course of 30 days in culture. As observed under basal conditions, the loss of *SOD1*$^{+/A4V}$ motor neurons was

greater than that of *SOD1*$^{+/+}$ control motor neurons. Two weeks of treatment with retigabine (1 μM) increased the number of ALS motor neurons *in vitro* by 25% ($p<10^{-4}$, t-test; Figure 2.13D), restoring the number of surviving ALS motor neurons to that of controls.

Inherent ER Stress in Human Motor Neurons is Dependent on Their Physiological Activity. Given that the *SOD1*$^{+/A4V}$ human MNs that we report here are hyper-excitable in comparison to controls and levels of *XBP1* splicing increased as MNs matured (Figure 2.11A), we reasoned that a relationship might exist between the inherent ER stress we found in MNs and their electrophysiological activity. Treatment of MN cultures with sufficient tetrodotoxin (TTX) to effectively block action potentials, as measured by MEA recording (Figure 2.14A), significantly reduced *XBP1* splicing (Figure 2.14B). Reciprocally, treatment of cultures with the glutamatergic agonist kainate, which depolarizes MNs, led to a significant increase in *XBP1* splicing (Figure 2.14D). Linopiridine, a compound that blocks Kv7 voltage gated potassium channels[59] and increased MN activity (Figure 2.14C), had a similar effect to kainate, in increasing *XBP1* splicing (Figure 2.14D). Thus, changes in the physiological activity of human MN cultures correlated with their levels of ER stress, with decreased or increased firing of action potentials leading to lower or higher ER stress levels, respectively. Consistent with these findings, in *SOD1*A4V MNs we also observed significant alterations in the transcription of *SCNA1* and *KCNN3,* which encode voltage-gated sodium and potassium channels, which could either be contributing or compensating for the changes in motor neuron excitability.

The ER plays an essential role in both neuronal proteostasis and physiology[62] and therefore we next assessed whether manipulating the level of ER stress would affect the electrical activity of MNs. Reduction of ER stress levels by treatment with salubrinal, resulted in a relative reduction in the number of spikes per minute (Figure 2.14E). Conversely, an acute treatment of MN cultures with DTT, which robustly induced *XBP1* splicing (Figure 2.14G), resulted in an increase in the number of spikes per minute as measured by multi-electrode array

recordings (Figure 2.14F). These two data sets suggest that ER stress, the UPR and the excitability of human MNs are interconnected and can potentially form a 'vicious' cycle driven by mutant SOD1 that may lead to apoptosis (Figure 2.14H).

Figure 2.14

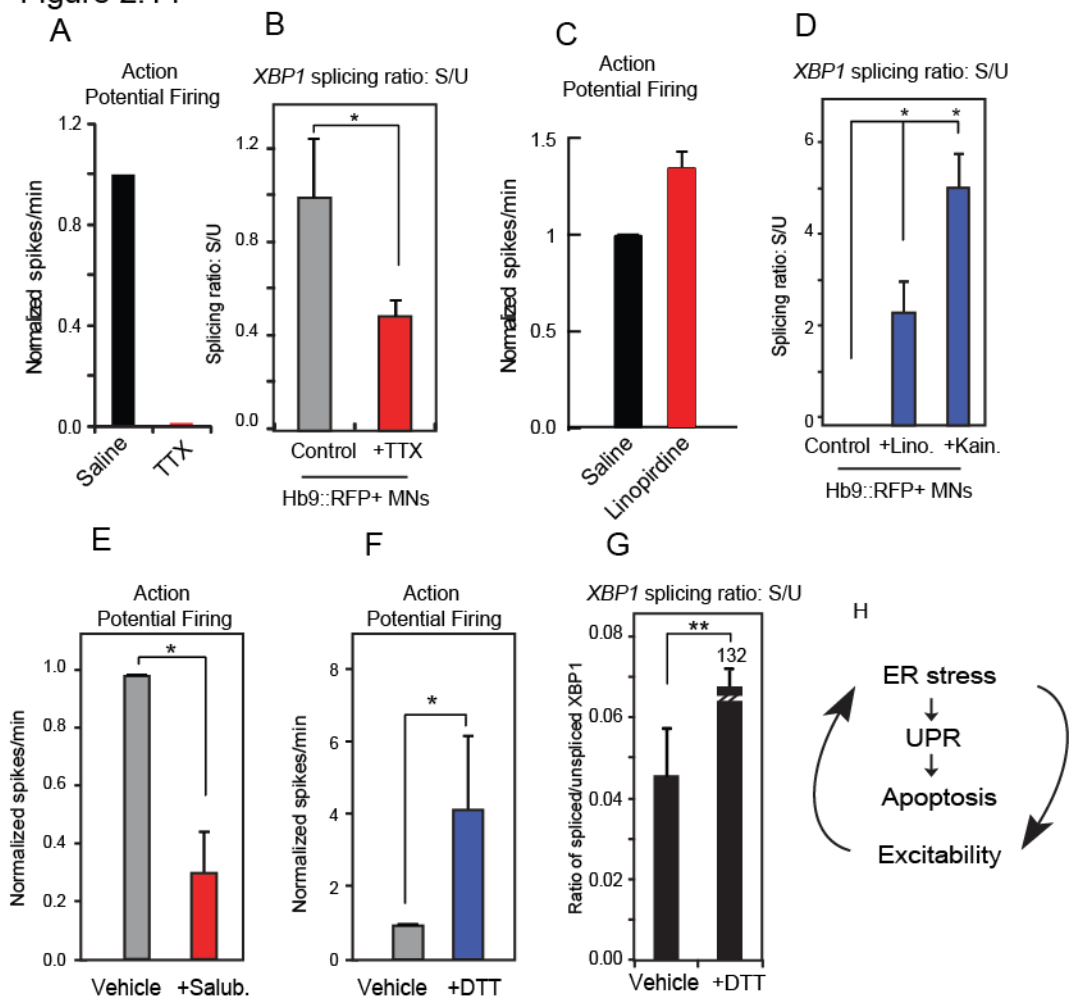

Figure 2.14. ER Stress In Human Motor Neurons Is Linked to Their Physiological Activity (A) TTX reduces the Electrophysiological activity and decreases XPB1 splicing (B). Conversly, Linopirdine increases action potential firing (C) and increases XPB1 splicing (D). Kainate also increases XPB1 splicing. Treatment with (E) salubrinal reduces, while (F) DTT increases action potential firing (n = 3, ±SEM, $p < 0.05$). DTT also increases XPB1 splicing (G). (H) ER stress, the UPR, and electrical activity of MNs are connected.

A Subset of Transcriptional Changes in *SOD1*$^{+/A4V}$ are shared in *C9orf72* MNs. We have shown that MNs produced from patient-derived iPSCs harboring the *SOD1*$^{+/A4V}$ genotype exhibit transcriptional changes and disturbances in mitochondrial metabolism and transport as well as oxidative and ER-related stress (Figures 2.9-2.11). Furthermore we demonstrated the

increased excitability of MNs generated from iPSCs harboring SOD1, FUS, and C9orf72 mutations. This finding begins to address the central question of whether mutations in the diverse genes that cause ALS converge on shared molecular pathways. To additionally focus on this question, we used two iPSC lines (19, RB8) that carried *GGGGCC* repeat expansions in the *C9orf72* locus. Both iPSCs were qualified via a range of assays including the Scorecard assay (Figure 2.15A). The presence of the hexanucleotide repeat expansion was confirmed by repeat prime PCR in both the parental fibroblasts as well as in multiple passages of the resulting iPSC lines (Figure 2.15B). Having demonstrated the ability of these lines to effectively differentiate into ISL/TUJ1 and HB9/TUJ1 postmitotic MNs (Figure 2.15C), we next sought to address whether these MNs exhibited transcriptional disturbances similar to the ones that we identified in *SOD1*$^{+/A4V}$ MNs.

In order to reduce artifacts due to line variability we selected six iPSC lines originating from five different healthy individuals[17]. These control iPSCs were generated via the same reprogramming strategy, had an almost identical average age at biopsy to the *C9orf72* ALS patients (52.4 Vs 52.5), and were maintained under the same culture conditions. We then differentiated all iPSC lines into MNs and FACS-purified *Hb9::*RFP+ MNs from healthy and *C9orf72* cultures in multiple biological replicates (n=8 iPSC lines, n=1-5 per line), after 15 days of co-culture with glial cells.

Using qRT-PCR we interrogated a subset of transcripts pre-selected to be representative of pathways, or cellular functions, which we had found to be altered by the *SOD1*A4V mutation (Figure 2.15D). These transcripts were selected based on their centrality in the particular pathway and/or the magnitude of their change in transcript abundance. Interestingly, in *C9orf72* mutant MNs, we did not detect a significant change in the transcript levels of selected genes implicated in electron transport in mitochondria (*COX7A2, COX7C, ACOT2*) but we did detect a significant change in levels of the mitochondrial transporter *MTX3,* suggesting that as in *SOD1*A4V MNs, mitochondrial function may be affected in *C9orf72* MNs.

Further supporting this view, we found that as in *SOD1*$^{+/A4V}$ MNs, *C9orf72* MNs exhibited a significant induction of catalase (*CAT),* indicative of oxidative stress.

Figure 2.15

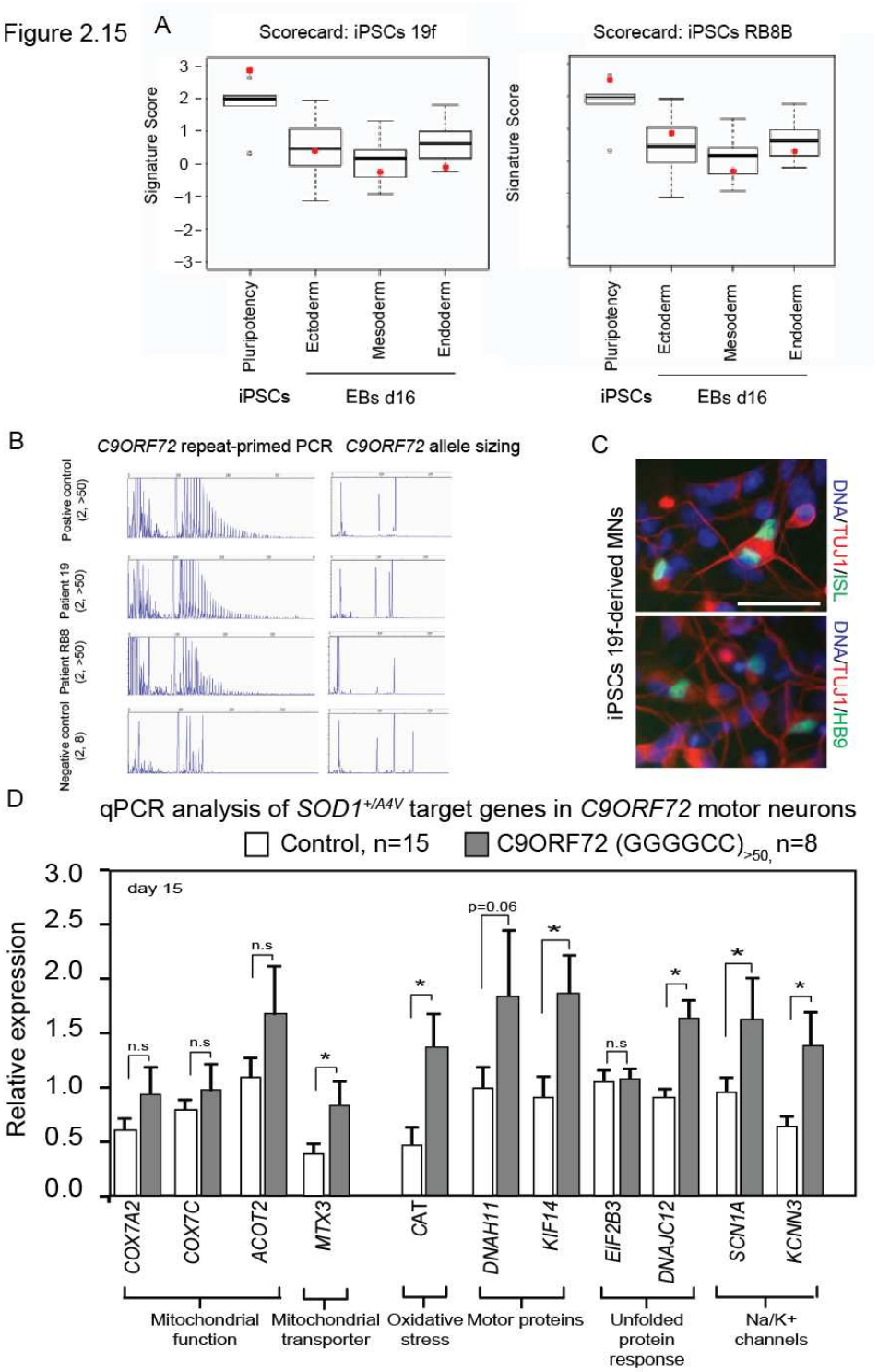

Figure 2.15. (Continued) Transcriptional Changes Detected in SOD1+/A4V MNs Are Partially Conserved in MNs with C9orf72 Expansions (A) Scorecard analysis of pluripotency genes for iPSCs and for three germ layer differentiation after 16 days, the red dots indicate values for each of the two lines. (B) Genotyping of samples by repeat-primed PCR and allele sizing for C9orf72. Patient samples RB8 and 19 exhibited more than 50 GGGGCC repeats. (C) Representative images of MN cultures differentiated from C9orf72 ALS patients (scale bar = 50 mm). (D) A subset of transcripts found to be differentially expressed in SOD1+/A4V MNs also exhibited significant transcriptional changes in MNs derived from C9orf72 (n = 2 lines, n = 8 replicates, ±SEM, $p < 0.05$) compared to a large number of control samples (n = 6 lines, n = 15 replicates, ±SEM, $p < 0.05$).

To determine whether intracellular transport might be impacted in *C9orf72* MNs, we examined expression of the two most highly upregulated motor proteins in *SOD1*$^{+/A4V}$ MNs (*KIF14* and *DNAH11*). As in *SOD1*$^{+/A4V}$ MNs, we found that expression of *KIF14* was significantly induced in *C9orf72* MNs. Although there was a trend towards increased *DNAH11* transcription, it did not reach significance.

Finally, we analyzed the levels of 4 additional transcripts (*DNAJC12, EIF2B3, SCN1A, KCNN3*) modulated in SOD1$^{+/A4V}$ MNs to determine whether there was evidence for an unfolded protein response and/or shared changes in the expression of channels that might be associated with changes in membrane excitability. Of these transcripts, 3 of 4, including both cation channels and the protein chaperone *DNAJC12* were transcribed at significantly different levels between control and *C9orf72* MNs, again supporting the notion that at least a subset of changes are shared between MNs of these two types.

Discussion

Transgenic rodent models of ALS have been indispensable for developing hypotheses concerning how mutant SOD1 proteins induce MN degeneration. However, studies in these animals have yet to yield an effective treatment[10,11]. We reasoned that the gulf between successful animal studies and more positive clinical outcomes might be bridged with model systems that enable hypotheses originating from animals to be validated and extended in the context of human MNs bearing patient mutations. Utilizing iPSCs, genome editing and directed differentiation, we established a well-controlled cell culture system and interrogated the differential properties of patient-derived and healthy control MNs. Advances in gene targeting tools, which we have employed here, allowed our studies of the *SOD1*A4V variant to be

elevated beyond correlative distinctions between ALS cases and controls[63] to the demonstration of causal connections between this mutation and transcriptional as well as functional phenotypes.

Notably, our studies demonstrate that the *SOD1*A4V missense mutation is necessary to cause a pro-apoptotic phenotype in cultured human MNs, restricting their long-term survival. By employing RNA sequencing, we defined the transcriptional differences between human *SOD1*$^{+/A4V}$ and isogenic control MNs. Curating these data supported the view that patient-specific ALS iPSC-derived MNs display hallmarks of disease found in both patients and in animal models. We identified defects in mitochondrial transport and morphology, oxidative and ER-related stress and an activated UPR, all of which were dependent on the presence of the *SOD1*A4V mutation. Importantly, we identified an increase in electrical activity in the mutant MNs which could be corrected by the FDA approved drug retigabine. The use of retigabine to block hyperexcitability of the *SOD1*$^{+/A4V}$ MNs rescued the survival deficit, implicating the altered physiology in the apoptotic phenotype described here.

We also found that other cell and neural-types were relatively unaffected by the *SOD1* mutation. However, and as is observed in ALS patients, our molecular and pharmacological studies suggest that human MNs were more susceptible to mutant *SOD1*. We propose that this susceptibility originates from a pre-existing burden of ER stress that we found to be constitutively present in normal physiologically active MNs, but absent from a variety of other cell-types. It has previously been proposed that combinations of stressors may converge and reinforce each other leading to dysfunction and eventual degeneration of vulnerable neurons[64]. We have found that the unfolded protein response, ER stress and electrical activity of MNs are interconnected and could therefore form a 'vicious cycle' driven by disease causing mutations such as *SOD1*A4V. Therefore three distinct categories of compounds may be of substantial therapeutic benefit to ALS patients by disrupting the 'vicious cycle'; those that support folding of proteins generally, those that specifically aid MNs in handling ER stress, and finally those that

reduce MN excitability. Our studies with salubrinal and retigabine, support this view. It is noteworthy that neither treatment with salubrinal, nor knockdown of *XBP1* alone resulted in a complete rescue of the survival deficit, implying that perhaps ER stress in only one of many components that contribute to MN death in our system.

While the function of the C9ORF72 protein remains unknown, the mechanism by which the hexanucleotide repeat expansion predisposes individuals to ALS has been suggested to range from haploinsufficiency to toxic gain-of-function properties of the mutant RNA or protein [6,65,66]. Our studies indicate that MNs derived from patients harboring the repeat expansion are hyper excitable and furthermore demonstrated a partial conservation of transcriptional changes between *SOD1*$^{+/A4V}$ and *C9orf72* cases. Of particular note are transcripts reflecting a heightened oxidative stress response, reduced mitochondrial function, as well as changes in cation channels and motor proteins. Our discoveries of transcriptional and functional aberrations particularly in relation to mitochondria and ER stress in these patient-specific iPSC MNs could potentially relate to the typically late clinical onset of ALS, as these pathways are known to be involved in natural ageing processes. Taken together our work validates the utility of iPSCs and genome engineering strategies for probing relationships between the genetic variants responsible for ALS in the MNs that selectively degenerate in this disease.

Experimental Procedures

Cell Culture. Cell cultures were maintained at 37^{C}, 5% CO^{2}. Human fibroblasts were cultured in KO-DMEM (Invitrogen), supplemented with 20% Earl's salts 199 (Gibco), 10% Hyclone (Gibco), 1x GlutaMax, (Invitrogen), and 100μM 2-mercaptoethanol and passaged by trypsinization (0.25% trypsin EDTA, Invitrogen). HuES and iPS cells were maintained on Matrigel (BD Biosciences) with mTeSR1 media (Stem Cell Technologies). Media was changed every 24 hours and lines were passaged by dispase (Gibco, 1mg/mL in hES media for 30min at 37^{C}). Human primary astrocytes were purchased from Sciencell Research Laboratories and

grown according to supplier instructions.

Derivation of Human Fibroblasts and iPS Cell Generation. Human fibroblasts were generated from 3mm forearm dermal biopsies following informed consent as described previously[14]. Generation of iPS cells was done essentially as reported previously by retroviral transduction of *KLF4*, *SOX2*, *OCT4* and *(cMYC)*[17]. Retrovirus preparations were done at the Harvard Gene Therapy Initiative at Harvard Medical School (Boston, MA). For iPS cell derivation, 30,000 human fibroblasts were transduced at an MOI of 3-5 per gene, with viruses containing all three genes.

Scorecard Assay. RNA samples were produced and analyzed as described previously (Bock et al., 2011) with minor modifications. Whole iPS and ES cell colonies were isolated by dispase treatment and plated in suspension in the presence of mTeSR1 media, cell aggregates (EBs) were allowed to form and 48hrs later EBs were switched to KOSR media without FGF (DMEM/F12, 10% KOSR, NEAA, Glutamax, and 100μM 2-mercaptoethanol). EBs were grown for a total of 16 days, at the end of which, cells were lysed and total RNA was extracted using Trizol (Invitrogen). Subsequently, RNA was analyzed on the NanoString nCounter using a custom codeset. The calculation of the iPS lines' lineage scores were performed according to our previously published protocol (Bock et al., 2011) using our published dataset for 20 human embryonic stem cell lines as a reference.

Motor Neuron Differentiation. MN differentiation was carried out as previously described[17] with modifications (see also Figure 2.2A). Briefly, pluripotent stem cell colonies were dissociated with accutase and single cells were plated in suspension in low-adherence dishes, at a 400K/ml density with 10mM ROCK inhibitor (Sigma, Y-27632) in mTeSR media for 24hrs. Embryoid bodies (EBs) were formed and media was gradually diluted (50% on day 3 and 100% on day 4) to KOSR (DMEM/F12, 10% KOSR) between days 1-4 and to a neural induction medium (NIM: DMEM/F12 with L-glutamine, NEAA, Heparin (2mg/ml), N2 supplement (Invitrogen) for days 5-24. From days 1-6 cells were cultured in the presence of SB431542

(10M, Sigma Aldrich) and Dorsmorphin (1mM, Stemgent), and from days 5-24 with BDNF (10mg/ml, R&D), ascorbic acid (AA, 0.4mg/ml, Sigma), Retinoic Acid (RA, 0.1mM, Sigma) and Smoothened Agonist 1.3 (SAG 1.3, 1mM, Calbiochem). At day 24 EBs were dissociated to single cells with Papain/DNase (Worthington Bio) and plated onto poly-lysine laminin-coated chamber slides/plates/coverslips (BD Biosciences) for relevant experiments.

Motor Neuron Survival Assay. After 24 days of differentiation, neuronal EBs were dissociated and 20K were plated on poly-D-lysine/Laminin coated 8-well chamber slides (BD biosciences) containing a confluent monolayer of primary cortical mouse glia. Primary glial preparations from P0-P2 mouse pups were generated as described previously[17]. Fresh glial preparations (<1 month, <2 passages) were used. Co-cultures were maintained in Neurobasal media (NB, Invitrogen), supplemented with B27 and N2 supplement (Invitrogen), 10mg/mL each of BDNF, GDNF, CNTF (R&D) and ascorbic acid (AA, 0.4mg/ml, Sigma) and fed every 2-3 days. Slides were fixed at various time points (3, 15, 20, 30 days), cultures were stained and cell numbers assessed. Whole-well images were quantified in a manner blinded to the genotype and condition of the experiment. Neuronal numbers on day 3 were set as 100% and numbers on subsequent time points were expressed as a percentage of day 3. To evaluate cell death, neuronal cultures were plated without glia on coverslips and live cells were assayed using the In Situ Cell Death Kit (Roche Diagnostics) according to manufacturer's instructions.

RNA Preparation, RT-PCR and RNA Sequencing. Total RNA was isolated from relevant cell types using Trizol LS (Invitrogen) according to manufacturer's instructions. MNs were differentiated as described above, MN cultures were plated on glial monolayers, infected with the Hb9::RFP virus on days 5-7 and FACS was used for purification of RFP MNs on day 15. A total of 300-1000ng was used to synthesize cDNA by reverse transcription according to the iSCRIPT kit (Bio-rad).Quantitative RT-PCR was then performed using SYBR green (Bio-Rad) and the iCycler system (Bio-rad). Quantitative levels for all genes were normalized to the average levels of 3 housekeeping genes: GAPDH/b-Actin/YWHAZ and expressed relative to the

relevant control samples or the lowest expressing sample in the experiment (see figure legends). QPCR for retroviral and endogenous reprogramming genes was carried out as previously reported[17]. All primer sequences are available upon request. Human spinal cord RNA was purchased from Clonetech (#636530) and total human brain RNA was purchased from Applied Biosystems (#AM6050). Spinal cord RNA was pooled from 22 male/female Caucasians aged between 22-69, while total brain RNA was pooled from 23 male/female Caucasians aged between 23-86. For TTX (48hrs), kainate/linopiridine (48hrs) and DTT (2hrs) treatments, co-cultures were treated, RFP positive neurons were purified and RNA was subsequently isolated. For next-generation RNA sequencing, RNA integrity numbers (RIN) above 7.5, determined by bioAnalyzer, were used for library preparation. In brief, RNA sequencing libraries were generated from ~250ng total RNA using the illumina TruSeq RNA kit v2, according to the manufacturer's directions. Libraries were sequenced at the Harvard Bauer Core Sequencing facility on a HiSeq 2000. All FASTQ files were analyzed using FastQC software (v 0.10.1) to confirm that Phred scores were acceptable at all read positions (median Phred score>25 and lower quartile>20). The FASTQ files were aligned to the GRCh37/hg19 reference genome using Tophat (v 2.0.7). Duplicated reads were removed using Picard Tools MarkDuplicates (v 1.44). Differential expression testing was performed independently using two separate analysis packages: Cufflinks (v 2.1.1) and DESeq. The Cufflinks output was visualized with the cummeRbund R package using a false discovery rate of 0.05. For DESeq analysis, gene level annotation count files were first generated using the HTseq count Python script (v 0.5.4). DESeq analysis was performed using the methods recommended by the package authors. Gene Ontology term enrichment was determined for significantly differentially expressed genes at a false discovery rate cutoff of 0.05 using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7. Gene Set Enrichment Analysis (GSEA, Broad Institute) was performed by first creating a pre-ranked gene list of all genes included in differential expression testing ordered by log2 fold change. Analysis was performed using the GSEA preranked tool

with the REACTOME and KEGG Pathway MSigDB collections.

Mitochondrial Transport Assays and EM Analysis. After 24 days of differentiation, neuronal EBs were dissociated and 20K were plated on poly-D-lysine/laminin-coated 35 mm glass bottom culture dishes (MatTek Corporation) or coverslips (BD-Biosciences). Cultures were infected with *Hb9::*RFP lentivirus 5 days after dissociation and MNs were selected based on expression of RFP. On days 23-26, MNs were stained with MitoTracker® Green FM (50nM, Invitrogen) and transferred to a custom observation chamber mounted on the stage of the microscope. Live microscopy of mitochondrial transport was performed with a Nikon Eclipse Ti equipped with an automated stage and In Vivo Scientific incubator. Mitochondrial movements were recorded for 5 minutes with 4-second time-lapse intervals using NIS-Elements (Nikon) using a 63x lens. Kymographs were generated from each video using NIS-Elements Analyzing Software (Nikon). Mitochondria were considered motile if they traveled faster than 0.017 µm/second. The average distance between mitochondria was calculated excluding motile mitochondria, and the total unoccupied space was divided by total process length analyzed to yield the proportion of processes unoccupied by mitochondria. Average distance was measured using the top portion of each kymograph using NIS-Elements Analyzing Software (Nikon). For Electron Microscopy analysis, MN cultures were fixed with 2.5% glutaraldehyde-2% paraformaldehyde in 0.1M sodium cacodylate buffer (pH 7.4) and maintained at 4℃ O/N. Cultures were then postfixed in 1% OsO4-1.5% KFeCN6 for 30 min, washed in water 3x and incubated in 1% aqueous uranyl acetate for 30mn followed by 2x washes in water and subsequent dehydration in grades of alcohol (5min each; 50%, 70%, 95%, 2x 100%). Cells were then embedded in plastic and ~60nm thick sections were cut, picked up onto copper grids, stained with lead citrate and analyzed in a JEOL 1200EX Transmission Electron Microscope. At least 3 independent differentiation experiments were analyzed in each case and pictures were taken by a technician blinded for sample IDs.

XBP1 Splicing Assay. 300ng of RNA was used to generate cDNA. PCR reactions were

set up using 2ml of cDNA and premixed *Ampligold Taq* Polymerase (Applied. Biosystems), at 66^{C} annealing. PCR products were analyzed on a 2% low-melting agarose gel. The ratio of spliced/unspliced bands was quantified using Image J software. Primers available upon request.

Gene Targeting. Zinc finger nucleases (ZFNs) were constructed using either the OPEN method as described previously[67] or a modified version of OPEN that uses antibiotic resistance for the selection as previously described[68]. Briefly, pools of ZF pre-selected zinc finger (ZF) domains were ligated together to create a combinatorial library of three-3 finger proteins. A bacterial two-hybrid-based selection system was used to interrogate the ZF library for proteins that could bind the appropriate target sequences of interest. ZF proteins that bound the target sequence were cloned into a mammalian expression vector and fused to heterodimeric FokI nuclease protein domains to construct ZFNs. Active ZFNs capable of inducing a double strand break at the desired locus were identified by screening pairs of nucleases for the capability to induce characteristic indel mutations at the SOD1 target site in sequencing the locus of ZFN-treated HEK 293s. 2.5 million iPS cells of the 39b cell line were accutased and nucleofected, using Human Stem Cell Kit II and program A-023, with 1 mg of ZFN plasmid and 5mg of targeting plasmid. After nucleofection the cells were plated on matrigel with mTesr and ROCK inhibitor. After 48hrs puromycin selection was applied for 1 week after which surviving colonies were passaged and gDNA was extractedPCR was used to confirm proper targeting of the cassette. To remove the puromycin cassette 2.5 million cells were nucleofected with 1mg of a mammalian expression plasmid containing hygromycin and 5mg of a mammalian expression plasmid containing the FLPo recombinase and plated on matrigel with mTeSR and ROCK inhibitor. Twenty four hours after nucleofection, hygromycin was added for 48 hours. Colonies were allowed to expand for 1 week then picked and genomic DNA was extracted. Sequencing of the genomic DNA was used to confirm removal of the puromycin cassette. SOD1 expression was verified by qPCR after RNA extraction and cDNA synthesis. PshAI digestion along with sequencing of the qPCR product demonstrated loss of expression of the mutant allele. Copy

number qPCR using primer SOD1cnF and SOD1cnR was performed as described previously[69] to rule out random integration events. Primer sequences are available upon request.

Genome Sequencing and Analysis. DNA samples were derived from the parental 39b cell line and the gene corrected clone using phenol chloroform extraction. The sequencing libraries were made with 50ng genomic DNA using the Illumina Nexterra DNA kit. Deep (30×) WGS was performed using the Illumina HiSeq 2500 Platform (500 bp library, 101 bp reads). All subsequent alignments and analysis were performed with hg19 as a reference. To investigate whether there were changes in copy number due to the cell line transformation, we used Genome STRiP[34] to extract and process the read depth signal from the aligned sequencing data. Using a genome alignability mask of size 101, we segmented the genome into non-overlapping windows each containing 100Kb of uniquely alignable base positions (based on a read length of 101bp). In each window, we computed normalized read depth by counting the read fragments aligned within that window with a minimum mapping quality of 10 and normalized the counts based on genome-wide sequencing depth and correcting for sequencing coverage bias due to local G+C content. This was done separately for the parental and derived cell lines. To look for regions of copy number change, we evaluated the ratio of normalized read depth in the derived cell line compared to the parental cell line in each window. To find rare coding SNPs in ALS genes, we annotated coding variants called by Haplotype caller with SNPeff[70]. SNPs classified as missense, silent, or nonsense were retained. We then integrated allele frequencies for the European population from the thousand genomes project[71]. Variants were selected that overlapped target genes for ALS. To find variants that differed between cell lines, we compared the genotypes of both lines in a stringent manner similar to the methodologies used to discover de novo mutations. For a variant to be confidently different between cell lines, we required a read depth of at least 2 and a likelihood score (PL) of at least 30 across both lines. Homozygous variants were required to have no more than 5% of the reads observed from the alternate allele, while heterozygous variants were required to have at least

30% of reads observed from the less frequent allele and at most 70% of the reads from the more frequent allele. To examine the off target effects of the designed nuclease, variants within the top 12,000 potential off target nuclease cut sites were selected from this filtered set of confident variants.

Nanostring Karyotyping. Karyotyping was undertaken using the Nanostring nCounter Human Karyotype Panel (Nanostring Technologies, USA) and performed as per the manufacturer's instructions. In brief, the protocol is as follows: 600 ng of genomic DNA was Alu1 digested at 37 ℃ for 2 hours, before being denatured at 95 ℃ for 5 minutes. To prevent renaturing samples were kept on ice. A total of 300 ng of Alu1-digested DNA per sample was mixed with hybridization buffer, capture and reporter codes. Following a 16 hour incubation at 65 ℃, samples were transferred to a Nanostring Prep station where hybridized DNA was bound to an imaging cartridge before imaging. Using reference samples, a copy number was calculated for each chromosome following normalization of the data using nSolver (Nanostring Technologies, USA) and Microsoft Excel.

MEA Recording. After 24 days of differentiation, equal numbers of control and ALS neurons were plated on poly-D-lysine/laminin coated p515A probes (Alpha Med Scientific) or M768-GLx 12-well plates (Axion BioSystems) at typical densities of 40,000-80,000/ probe or well. Adherence to the electrodes after plating over the course of subsequent incubation was variable, with occasional migration of neurons off the probes. For that reason, all probes were visualized immediately before each recording session to confirm a full monolayer of cells. Initial experiments (11a, 18a, 39b, and RB9d comparison) were performed as close as possible to the time of patch recordings (4 weeks). However, because we wished to evaluate firing at a time point prior to significant motor neuron death (Kiskinis et al., cosubmission), we performed subsequent experiments (39b-Cor and 39b comparison and all later experiments) at day 14 after dissociation.

Recordings from 64 extracellular electrodes were made using a Med64 (Alpha Med

Scientific) or Maestro (Axion BioSystems) MEA recording amplifier with a head stage that maintained a temperature of 37℃. For Med64 recordi ngs, data were sampled at 20 kHz, digitized, and analyzed by spike clustering and spike extraction algorithms using Mobius software (Alpha Med Scientific) with a 2 kHz 9-pole Bessel low pass filter, 10 µV action threshold detection limit, and 30% cluster similarity radius. These standard settings were maintained for all analyses. We confirmed that we obtained similar results across a wide range of action potential threshold and cluster similarity radius settings. Cross-correlation analysis to detect and correct for clusters detected by multiple electrodes was performed using custom Matlab software. Total action potential firing rates and mean neuronal firing frequencies were then determined and plotted. In order to record in larger replicates, we used the Axion Maestro MEA device, in a 12-well format with 64 extracellular electrodes in each well. For Maestro recordings, data were sampled at 15 kHz, digitized, and analyzed using Axion Integrated Studio software (Axion BioSystems) with a 200 Hz high pass and 2500 kHz low pass filter and an adaptive spike detection threshold set at 5.5 times the standard deviation for each electrode with 1 second binning. These standard settings were maintained for all Axion MEA recording and analysis.

For retigabine dose response curves, action potential numbers during one minute of recording in each concentration of retigabine were normalized to the initial action potential number during one minute of recording in standard extracellular saline solution. The EC50 value was determined by fitting the mean normalized data values to the Hill equation, $y = 1/((EC50/x)^{nH}+1)$ where nH is the Hill coefficient.

Patch Electrophysiology. Whole-cell patch recordings were performed on iPS-derived motor neurons identified by transduction with an *HB9*::RFP lentivirus. Lentiviral transduction was typically performed 7-10 days before recording. Two large comparisons were performed, one consisting of 11a, 18a, 39b, and RB9d, and the second consisting of 39b-Cor and 39b. Each comparison was made from pooled data from multiple separate experiments, each

consisting of synchronous and parallel iPSC culture and differentiation, embryoid body dissociation, plating and maturation of control- and ALS-derived neurons. Equal numbers of control and ALS motor neurons were recorded from each experiment. Comparison of the original four lines (11a, 18a, 39b, and RB9d) was made using four separate parallel differentiation experiments, while comparison of the isogenic correction comparison (39b-Cor and 39b) was performed using three separate parallel differentiation experiments.

For each experiment, neurons from control and ALS lines were dissociated after 24 days of differentiation and plated onto poly-d-lysine/laminin coated glass coverslips (20,000-40,000/coverslip) and allowed to mature for four weeks from start of differentiation. We chose four weeks as the best timepoint because this yielded the most homogeneous population of mature-appearing *HB9*::RFP-positive motor neurons (at the requisite low cell density for patch clamp) with the most mature electrophysiological properties. Whole-cell current-clamp and voltage-clamp recordings were performed using a Multiclamp 700B (Molecular Devices) at room temperature (21-23℃). Data were sampled at 20 kHz and digitized with a Digidata 1440A A/D interface and recorded using pCLAMP 10 software (Molecular Devices). Data were low-pass filtered at 2 kHz. Patch pipettes were pulled from borosilicate glass capillaries on a Sutter Instruments P-97 puller and had resistances of 2-4 MW. The pipette capacitance was reduced by wrapping the shank with Parafilm and compensated for using the amplifier circuitry. Series resistance was typically 5-10 MW, always less than 15 MW, and compensated by at least 80%. Neurons were excluded from analysis if holding current at -80 mV exceeded 100 pA, input resistance was less than 250 or greater than 2000 MW, or spikes elicited from -65 mV had peaks below 0 mV. Resting membrane potential was determined by averaging for 20s of recording, and afterwards a small holding current (typically with amplitude less than 5 pA) was used to clamp the resting membrane potential as close as possible to -65 mV. Rheobase was measured by applying 1 s steps in increments of 2.5 pA until an action potential was generated. Current ramps were elicited from an initial hyperpolarizing current of 10 pA for 1 s followed by a

210 pA/s depolarizing ramp of duration 1 s. Spikes on the ramps were counted if the peak voltage exceeded -10 mV. Action potential properties were determined using custom-written analysis software in Igor Pro (Wavemetrics) with DataAccess (Bruxton) for importing the files. For voltage-clamp recordings, voltages were elicited by 100-ms depolarizing steps from a holding potential of -80 mV to test potentials ranging from -80 mV to 50 mV in 10 mV increments. For the latter gene correction experiments the step length was increased to 200 ms to assay delayed-rectifier currents after more complete decay of transient potassium currents. For retigabine patch applications, resting membrane potential was recorded immediately before and 10 seconds after the application of 10 µM retigabine (in each case, membrane potential was an average of values sampled for 20 seconds). For all patch experiments, series resistance was monitored by brief -5 pA hyperpolarizing steps during current clamp recordings and by 5 mV hyperpolarizing steps during voltage clamp recordings. Electrode drift was measured at the end of each recording and was typically 1-2 mV. The extracellular solution was sodium-based and contained 135 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM glucose, 10 mM HEPES 10, pH 7.4. The intracellular solution was potassium-based and contained 150 mM KCl, 2 mM $MgCl_2$, 10 mM HEPES, 4 mM MgATP, 0.3 mM NaGTP, 10 mM Na_2PhosCr, 1mM EGTA, pH 7.4. For isolation of delayed-rectifier potassium channels, 300 nM TTX and 100 µM $CdCl_2$ were used to block voltage-gated sodium channels and voltage- and calcium-activated potassium channels, respectively. For isolation of voltage-gated sodium currents, internal KCl was replaced by CsCl to block potassium currents and 100 µM $CdCl_2$ was used to block calcium currents.

Treatments with Small Molecules and siRNAs. For ER stress induction, human fibroblasts and human astrocytes were treated with Dithiothreitol (DTT) (2mM) (Bio-Rad Laboratories) for 2hrs washed with PBS and either fixed, stained and evaluated or trypsinized for RNA/protein collection. For analysis of SOD1 protein, MN cultures were treated with vehicle or 1mM MG132 for 48hrs. For assessing the role of XBP1 (s14913) and ATF4 (s1702) 40nM of

siRNA (Ambion) was transfected into MN cultures using siRNA-Select in Optimem on days 10/20 and knockdown levels and survival were assessed on day 30.

Immunocytochemistry. Cell cultures were fixed in 4% PFA for 15minutes at 4℃, permeabilized with 0.2% Triton-X in PBS for 45 minutes and blocked with 10% donkey serum in PBS-T (Triton 0.1%). Cells were then incubated in primary antibody overnight and secondary antibodies for 1 hour in 2% donkey serum in PBS-T after several washes in between. DNA was visualized by a Hoechst stain. The following antibodies were used: Primary antibodies used in this study are TRA1-81 (1:500, Chemicon, MAB4381), Nanog (1:500, R&D, AF 1997), Islet1 (1:200, DSHB, 40.2D6), HB9 (1:100, DSHB, MNR2 81.5C10-c), ChAT (1:100, Chemicon, AB144P), TUJ1 (1:1000, Sigma, T2200), MAP2 (1:10000, Abcam ab5392), BrdU (3H579, Santa Cruz Biotechnology, sc-70441), Ki67 (1:400, Abcam, ab833), GFP (1:500, Life Technologies, A10262), SOD1 (1:2000, Agrisera #AS09 540), Hoxa5 and FOXP1 (courtesy of Susan Morton, Jessell lab). Secondary antibodies used (488, 555, 594, and 647) were AlexaFluor (1:1000, Life Technologies) and DyLight (1:500, Jackson ImmunoResearch Laboratories).

Chick Embryo Transplants. HUES3 Hb9::GFP+ MNs were differentiated and one EB was placed into the neural tube of a stage 15 chick embryo using a tungsten needle. Embryos were harvested after 5 days, fixed and sectioned in paraffin. Sections were immunostained with antibodies specific for GFP (rabbit, Abcam) and an HRP-conjugated anti-rabbit antibody (Vector labs) and counterstained with hematoxylin.

Western Blots. For analysis of Phospho-eIF2a protein, cells were lysed in RIPA buffer (150mM Sodium Chloride; 1% Triton X-100; 0.5% sodium deoxycholate; 0.1% SDS; 50 mM Tris pH 8.0) containing protease and phosphatase inhibitors (Roche) for 20 min on ice, and centrifuged at high speed. Samples containing 20mg of protein were separated by SDS-PAGE (NuPAGE®, Life Technologies) and transferred to nitrocellulose membranes. Membranes were probed with anti-Phospho-eIF2a antibody (#3597, Cell Signaling Technology), and anti-a-

Tubulin (abcam, ab4074) and anti-eIF2a (Cell Signaling Technology, #9722) antibodies were used as loading and normalizing controls, respectively. For analysis of SOD1 protein, detergent-soluble fractions were prepared using RIPA buffer and detergent-insoluble fractions were obtained using UREA buffer (8M UREA; 4% CHAPS; 40 mM Tris; 0.2% Bio-Lyte® 3/10 ampholyte). 5mg of detergent-soluble and equivalent volumes of detergent-insoluble protein samples were separated by SDS-PAGE (Bio Rad Laboratories), transferred to PDVF membranes and probed with anti-SOD1 antibody (Agrisera #AS09 540) and anti-a-Tubulin (Sigma Aldrich # T6199). For mitochondrial biogenesis analysis, 6ug of detergent soluble protein samples were analyzed using the MitoBiogenesis^{TM} Western Blot Cocktail (ab123545).

Statistical analysis. Statistical significance was assessed by a standard Students T test (1 tail & 2 tail); $P<0.05$ was considered significant. Two-tailed, unpaired tests were used except to confirm specific hypotheses, in which case one-tailed, unpaired tests were used. Error bars represent ±s.e.m, unless otherwise stated.

References

1 Cleveland, D. W. & Rothstein, J. D. From Charcot to Lou Gehrig: deciphering selective motor neuron death in ALS. *Nat Rev Neurosci* **2**, 806-819, doi:10.1038/35097565 (2001).

2 Hardiman, O., van den Berg, L. H. & Kiernan, M. C. Clinical diagnosis and management of amyotrophic lateral sclerosis. *Nat Rev Neurol* **7**, 639-649, doi:10.1038/nrneurol.2011.153 (2011).

3 Al-Chalabi, A. *et al.* The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta Neuropathol* **124**, 339-352, doi:10.1007/s00401-012-1022-4 (2012).

4 Sreedharan, J. & Brown, R. H., Jr. Amyotrophic lateral sclerosis: Problems and prospects. *Ann Neurol*, doi:10.1002/ana.24012 (2013).

5 Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59-62 (1993).

6 Renton, A. E. *et al.* A Hexanucleotide Repeat Expansion in< i> C9ORF72</i> Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **72**, 257-268 (2011).

7 Gurney, M. E. *et al.* Motor neuron degeneration in mice that express a human Cu, Zn superoxide dismutase mutation. *SCIENCE-NEW YORK THEN WASHINGTON-*, 1772-1772 (1994).

8 Howland, D. S. *et al.* Focal loss of the glutamate transporter EAAT2 in a transgenic rat model of SOD1 mutant-mediated amyotrophic lateral sclerosis (ALS). *Proc Natl Acad Sci U S A* **99**, 1604-1609, doi:10.1073/pnas.032539299 (2002).

9 Oh, Y. K., Shin, K. S., Yuan, J. & Kang, S. J. Superoxide dismutase 1 mutants related to amyotrophic lateral sclerosis induce endoplasmic stress in neuro2a cells. *J Neurochem* **104**, 993-1005, doi:10.1111/j.1471-4159.2007.05053.x (2008).

10 Benatar, M. Lost in translation: treatment trials in the SOD1 mouse and in human ALS. *Neurobiol Dis* **26**, 1-13, doi:10.1016/j.nbd.2006.12.015 (2007).

11 Gladman, M., Cudkowicz, M. & Zinman, L. Enhancing clinical trials in neurodegenerative disorders: lessons from amyotrophic lateral sclerosis. *Curr Opin Neurol* **25**, 735-742, doi:10.1097/WCO.0b013e32835a309d (2012).

12 Sreedharan, J. *et al.* TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* **319**, 1668-1672, doi:10.1126/science.1154584 (2008).

13 Ling, S. C., Polymenidou, M. & Cleveland, D. W. Converging Mechanisms in ALS and FTD: Disrupted RNA and Protein Homeostasis. *Neuron* **79**, 416-438, doi:10.1016/j.neuron.2013.07.033 (2013).

14 Dimos, J. *et al.* Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science (New York, N.Y.)* **321**, 1218-1221, doi:10.1126/science.1158799 (2008).

15 Kiskinis, E. & Eggan, K. Progress toward the clinical application of patient-specific pluripotent stem cells. *J Clin Invest* **120**, 51-59, doi:10.1172/jci40553 (2010).

16 Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872 (2007).

17 Boulting, G. L. *et al.* A functionally characterized test set of human induced pluripotent stem cells. *Nature biotechnology* **29**, 279-286 (2011).

18 Wichterle, H., Lieberam, I. & Porter, J. Directed differentiation of embryonic stem cells into motor neurons. *Cell* (2002).

19 Bilican, B. *et al.* Mutant induced pluripotent stem cell lines recapitulate aspects of TDP-43 proteinopathies and reveal cell-specific vulnerability. *Proceedings of the National Academy of Sciences* **109**, 5803-5808 (2012).

20 Donnelly, C. J. *et al.* RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* **80**, 415-428, doi:10.1016/j.neuron.2013.10.015 (2013).

21 Egawa, N. *et al.* Drug screening for ALS using patient-specific induced pluripotent stem cells. *Sci Transl Med* **4**, 145ra104, doi:10.1126/scitranslmed.3004052 (2012).

22 Chambers, S. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature biotechnology* **27**, 275-280, doi:10.1038/nbt.1529 (2009).

23 Amoroso, M. W. *et al.* Accelerated High-Yield Generation of Limb-Innervating Motor Neurons from Human Stem Cells. *J Neurosci* **33**, 574-586, doi:10.1523/jneurosci.0906-12.2013 (2013).

24 Kanning, K. C., Kaplan, A. & Henderson, C. E. Motor neuron diversity in development and disease. *Annu Rev Neurosci* **33**, 409-440, doi:10.1146/annurev.neuro.051508.135722 (2010).

25 Marchetto, M. C. N. *et al.* Non-Cell-Autonomous Effect of Human SOD1< sup> G37R</sup> Astrocytes on Motor Neurons Derived from Human Embryonic Stem Cells. *Cell Stem Cell* **3**, 649-657 (2008).

26 Kiernan, J. A. & Hudson, A. J. Changes in sizes of cortical and lower motor neurons in amyotrophic lateral sclerosis. *Brain* **114 (Pt 2)**, 843-853 (1991).

27 McIlwain, D. L. Nuclear and cell body size in spinal motor neurons. *Adv Neurol* **56**, 67-74 (1991).

28 Di Giorgio, F. P., Boulting, G. L., Bobrowicz, S. & Eggan, K. C. Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. *Cell Stem Cell* **3**, 637-648, doi:S1934-5909(08)00522-5 [pii] 10.1016/j.stem.2008.09.017 [doi] (2008).

29 Bock, C. *et al.* Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439-452 (2011).

30 Mekhoubad, S. *et al.* Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* **10**, 595-609, doi:10.1016/j.stem.2012.02.014 (2012).

31 Lattante, S. *et al.* Contribution of major amyotrophic lateral sclerosis genes to the etiology of sporadic disease. *Neurology* **79**, 66-72, doi:10.1212/WNL.0b013e31825dceca (2012).

32 Cardoso, R. M. *et al.* Insights into Lou Gehrig's disease from the structure and instability of the A4V mutant of human Cu, Zn superoxide dismutase. *Journal of molecular biology* **324**, 247-256 (2002).

33 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297-1303, doi:10.1101/gr.107524.110 (2010).

34 Handsaker, R. E., Korn, J. M., Nemesh, J. & McCarroll, S. A. Discovery and genotyping of genome structural polymorphism by sequencing on a population scale. *Nat Genet* **43**, 269-276, doi:10.1038/ng.768 (2011).

35 Brotherton, T. E. *et al.* Localization of a toxic form of superoxide dismutase 1 protein to pathologically affected tissues in familial ALS. *Proc Natl Acad Sci U S A* **109**, 5505-5510, doi:10.1073/pnas.1115009109 (2012).

36 Bruijn, L. I. *et al.* Aggregation and motor neuron toxicity of an ALS-linked SOD1 mutant independent from wild-type SOD1. *Science* **281**, 1851-1854 (1998).

37 Shibata, N. *et al.* Intense superoxide dismutase-1 immunoreactivity in intracytoplasmic hyaline inclusions of familial amyotrophic lateral sclerosis with posterior column involvement. *J Neuropathol Exp Neurol* **55**, 481-490 (1996).

38 Bosco, D. *et al.* Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS. *Nature neuroscience* **13**, 1396-1403, doi:10.1038/nn.2660 (2010).

39 Deng, H.-X. *et al.* Conversion to the amyotrophic lateral sclerosis phenotype is associated with intermolecular linked insoluble aggregates of SOD1 in mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 7142-7147, doi:10.1073/pnas.0602046103 (2006).

40 Hunt, M. C., Solaas, K., Kase, B. F. & Alexson, S. E. Characterization of an acyl-coA thioesterase that functions as a major regulator of peroxisomal lipid metabolism. *J Biol Chem* **277**, 1128-1138, doi:10.1074/jbc.M106458200 (2002).

41 Krasnova, I. N. *et al.* CREB phosphorylation regulates striatal transcriptional responses in the self-administration model of methamphetamine addiction in the rat. *Neurobiol Dis* **58**, 132-143, doi:10.1016/j.nbd.2013.05.009 (2013).

42 Minkiewicz, J., de Rivero Vaccari, J. P. & Keane, R. W. Human astrocytes express a novel NLRP2 inflammasome. *Glia* **61**, 1113-1121, doi:10.1002/glia.22499 (2013).

43 Zoumakis, E., Rice, K. C., Gold, P. W. & Chrousos, G. P. Potential uses of corticotropin-releasing hormone antagonists. *Ann N Y Acad Sci* **1083**, 239-251, doi:10.1196/annals.1367.021 (2006).

44 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).

45 Mootha, V. K. *et al.* Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell* **115**, 629-640 (2003).

46 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

47 Trusina, A., Papa, F. R. & Tang, C. Rationalizing translation attenuation in the network architecture of the unfolded protein response. *Proc Natl Acad Sci U S A* **105**, 20280-20285, doi:10.1073/pnas.0803476105 (2008).

48 Hetz, C. *et al.* XBP-1 deficiency in the nervous system protects against amyotrophic lateral sclerosis by increasing autophagy. *Genes & development* **23**, 2294-2306, doi:10.1101/gad.1830709 (2009).

49 Ron, D. & Walter, P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol* **8**, 519-529, doi:10.1038/nrm2199 (2007).

50 Atkin, J. D. *et al.* Induction of the unfolded protein response in familial amyotrophic lateral sclerosis and association of protein-disulfide isomerase with superoxide dismutase 1. *J Biol Chem* **281**, 30152-30165, doi:10.1074/jbc.M603393200 (2006).

51 Atkin, J. D. *et al.* Endoplasmic reticulum stress and induction of the unfolded protein response in human sporadic amyotrophic lateral sclerosis. *Neurobiol Dis* **30**, 400-407, doi:10.1016/j.nbd.2008.02.009 (2008).

52 Saxena, S., Cabuy, E. & Caroni, P. A role for motoneuron subtype–selective ER stress in disease manifestations of FALS mice. *Nature neuroscience* **12**, 627-636 (2009).

53 Boyce, M. *et al.* A selective inhibitor of eIF2alpha dephosphorylation protects cells from ER stress. *Science* **307**, 935-939, doi:10.1126/science.1101902 (2005).

54 Bostock, H., Sharief, M. K., Reid, G. & Murray, N. M. Axonal ion channel dysfunction in amyotrophic lateral sclerosis. *Brain* **118 (Pt 1)**, 217-225 (1995).

55 Kanai, K. *et al.* Altered axonal excitability properties in amyotrophic lateral sclerosis: impaired potassium channel function related to disease stage. *Brain* **129**, 953-962, doi:10.1093/brain/awl024 (2006).

56 Kanai, K. *et al.* Motor axonal excitability properties are strong predictors for survival in amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry* **83**, 734-738, doi:10.1136/jnnp-2011-301782 (2012).

57 Tamura, N. *et al.* Increased nodal persistent Na+ currents in human neuropathy and motor neuron disease estimated by latent addition. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* **117**, 2451-2458, doi:10.1016/j.clinph.2006.07.309 (2006).

58 Fritz, E. *et al.* Mutant SOD1-expressing astrocytes release toxic factors that trigger motoneuron death by inducing hyperexcitability. *Journal of neurophysiology* **109**, 2803-2814 (2013).

59 Brown, D. A. & Passmore, G. M. Neural KCNQ (Kv7) channels. *British journal of pharmacology* **156**, 1185-1195 (2009).

60 Porter, R. J., Nohria, V. & Rundfeldt, C. Retigabine. *Neurotherapeutics* **4**, 149-154 (2007).

61 Wickenden, A. D., Yu, W., Zou, A., Jegla, T. & Wagoner, P. K. Retigabine, a novel anti-convulsant, enhances activation of KCNQ2/Q3 potassium channels. *Molecular pharmacology* **58**, 591-600 (2000).

62 Verkhratsky, A. Physiology and pathophysiology of the calcium store in the endoplasmic reticulum of neurons. *Physiol Rev* **85**, 201-279, doi:10.1152/physrev.00004.2004 (2005).

63 Soldner, F. & Jaenisch, R. Medicine. iPSC disease modeling. *Science* **338**, 1155-1156, doi:10.1126/science.1227682 (2012).

64 Saxena, S. & Caroni, P. Selective neuronal vulnerability in neurodegenerative diseases: from stressor thresholds to degeneration. *Neuron* **71**, 35-48, doi:10.1016/j.neuron.2011.06.031 (2011).

65 Ash, P. E. *et al.* Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77**, 639-646, doi:10.1016/j.neuron.2013.02.004 (2013).

66 Mori, K. *et al.* The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. *Science*, doi:10.1126/science.1232927 (2013).

67 Maeder, M. *et al.* Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Molecular cell* **31**, 294-301, doi:10.1016/j.molcel.2008.06.016 (2008).

68 Sander, J. *et al.* Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nature methods* **8**, 67-69, doi:10.1038/nmeth.1542 (2011).

69 D'Haene, B., Vandesompele, J. & Hellemans, J. Accurate and objective copy number profiling using real-time quantitative PCR. *Methods* **50**, 262-270, doi:10.1016/j.ymeth.2009.12.007 (2010).

70 Abecasis, G. R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65, doi:10.1038/nature11632 (2012).

71 Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80-92, doi:10.4161/fly.19695 (2012).

Chapter 3: Advancing the study of SOD1 A4V with purified MN cultures.

Introduction

For decades, the study and development of novel therapeutics for neurodegenerative diseases have been stymied by difficulties accessing the specific types of neuronal tissues which degenerate in these conditions[1]. Pluripotent stem cell technology holds great promise for ameliorating this problem, theoretically allowing for the production of an unlimited supply of the affected cell type, as well as patient-specific cell lines containing mutations which may predispose an individual toward the development of neurodegenerative pathology[1].

Amyotrophic Lateral Sclerosis (ALS) is the most common adult motor neuron (MN) neurodegenerative disorder[2]. Spinal and cortical MNs are the primary cell types which degenerate during ALS disease progression, leading to paralysis and eventually death[2]. To date, mutations in over 20 genes have been identified which can result in familial forms of ALS[2,3], the first of which was found in *SUPEROXIDE DISMUTASE 1 (SOD1). SOD1* is a free radical scavenger expressed widely throughout the body[3]. A mouse model overexpressing a mutated version of human *SOD1* displays many of the pathologies found in patients, and has been used extensively to understand the disorder[4]. This mouse model has been extremely valuable for studying ALS, resulting in several novel insights about the disease[5,6]. However, the mutant protein must be expressed well above normal physiological levels for phenotypes to develop[7]. Furthermore, drugs found to be efficacious in the animal model have failed in human clinical trials[8], suggesting an incomplete concordance between the mouse model and human patients, and motivating the development of ALS models using human cell lines. To address this concern, we and others have demonstrated the application of human pluripotent stem cell technology for modeling ALS, providing an in vitro human system in which to discover disease relevant phenotypes and test possible therapeutics[9].

Our earlier studies reported on the establishment of many in vitro phenotypes which were dependent on the *SOD1 A4V* mutation[9]. Specifically, we used gene targeting to correct the mutation, and demonstrated the requirement of mutated *SOD1* for the development of several in

vitro phenotypes including malformed mitochondria, reduced MN survival, increased neuronal excitability, and increased ER stress[9]. We wanted to delve deeper into the in vitro phenotypes we previously described by working toward improved understanding of two specific areas.

First, while gene correction was used to determine the necessity of the A4V mutation for the development of the phenotypes described in Chapter 2[9], a single cell line was corrected, leaving open the possibility that the patient-specific iPSC line used for gene correction had other unknown modifying mutations which led to more dramatic phenotypes in vitro. To determine which phenotypes may have broader penetrance and therefore may be more clinically relevant, we aimed to further probe the effect of the *SOD1 A4V* mutation on stem cell-derived MNs, specifically determining the sufficiency of this mutation for the development of the in vitro phenotypes by introducing the mutation into a control ES cell genetic background. Because the hES cell line was not derived from patient tissue with any clinical history of ALS, it should provide an unaffected control genetic background in which to induce the *SOD1 A4V* mutation[10].

Second, while the degeneration of MNs is the primary cause of patient death, other cells types have been implicated in ALS disease progression[11,12]. This non-cell autonomous aspect of the disorder can lead to complications when studying the disease with pluripotent stem cell technology, as current in vitro differentiation protocols produce a panoply of cell types in addition to the specific cell type of interest[13]. When purely cell-autonomous aspects of the disease are to be studied, it would be ideal to isolate the MNs away from any other neuronal byproducts which could be expressing the mutation being studied, and therefore could have a non-cell autonomous role in degeneration of the MNs[1]. To address this limitation, we used nuclease assisted gene targeting to induce the *SOD1 A4V* mutation in a previously described reporter hES cell line (Hues3 HB9:GFP)[14]. This cell line harbors a transgenic construct (HB9:GFP) enabling the transcription of GFP to be controlled by the motor neuron specific HB9 promotor[14]. Using gene targeting to introduce the A4V mutations into the Hues3 HB9:GFP cell line allowed for the purification of MNs from heterogeneous neuronal differentiations of isogenic

control and *SOD1*$^{A4V/WT}$ pluripotent cell lines.

The current chapter will describe the use of a combination of gene targeting and a reporter hES cell line to address the questions described above, and demonstrate the sufficiency of the *SOD1 A4V* variant to drive the accumulation of abnormal mitochondria and increased neuronal excitability. Furthermore, we used transcriptional profiling by RNA sequencing to probe the MN transcripts for proximal targets which may be responsible for the increase in neuronal excitability observed in MNs expressing the *SOD1 A4V* variant. Lastly, using transcriptional profiling, we made initial attempts at discovering specific target genes which, when modulated by siRNA, can result in increased neuronal excitability.

Results

Induction of *SOD1 A4V* mutation in control hES cell line. To introduce the *SOD1 A4V* mutation into the previously validated Hues3 HB9::GFP *SOD1*$^{WT/WT}$ reporter cell line[14], we used similar techniques and reagents as described previously in Chapter 2 of this thesis[9]. Briefly, a targeting plasmid was constructed containing a FRT flanked puromycin cassette surrounded by 700 bp homology arms for the *SOD1* locus (Figure 3.1A). This targeting plasmid was introduced into the HuES3 HB9:GFP cell line, along with a plasmid encoding a previously described zinc finger nuclease capable of introducing a DSB at the *SOD1* locus. The introduction of a DSB at the *SOD1* locus stimulated homologous recombination between the genomic and plasmid DNA, resulting in the incorporation of the *SOD1 A4V* variant along with the puromycin cassette under the control of the PGK promotor to identify successful targeting events by antibiotic selection. This gene targeting strategy resulted in the production of several puromycin-resistant stem cell colonies which were picked and allowed to expand. PCR followed by sequencing and restriction fragment length polymorphism (RFLP) analysis for a unique PshAI site were used to identify properly targeted colonies (Figure 3.1B and C). Two correctly targeted stem cell lines from independent targeting experiments were chosen for excision of the puromycin cassette which from here out will be described as Hues3 *SOD1*$^{KO/WT}$ ZFN_1 and

Hues3 *SOD1*$^{KO/WT}$ ZFN_2. Separate expression plasmids containing the site specific recombinase FLPo and Hygromycin expression cassettes were co-transfected into the positive stem cell lines and hygromycin was applied to select for cells which were successfully transfected. Resistant colonies were picked and expanded, and PCR followed by RFLP analysis was used to confirm the removal of the puromycin cassette (Figure 3.1B and C). Single colonies derived from Hues3 *SOD1*$^{KO/WT}$ ZFN_1 and Hues3 *SOD1*$^{KO/WT}$ ZFN_2 which had their puromycin cassettes removed were named Hues3 *SOD1*$^{A4V/WT}$ ZFN_1 and Hues3 *SOD1*$^{A4V/WT}$ ZFN_2 respectively. Because the puromycin cassette was integrated into the first intron of the *SOD1* locus during the first step of the gene targeting, transcription and translation of the full length *SOD1* mRNA and protein were inhibited in the *SOD1*$^{KO/WT}$ cell lines. Western blotting and qPCR of the parental Hues3 *SOD1*$^{WT/WT}$, Hues3 *SOD1*$^{KO/WT}$ ZFN_1, and Hues3 *SOD1*$^{A4V/WT}$ ZFN_1 confirmed the reduction in *SOD1* expression when the puromycin cassette was introduced into the *SOD1* locus and the subsequent re-establishment of *SOD1* transcript and protein levels after removal of the puromycin cassette (Figure 3.1D and E).

Figure 3.1

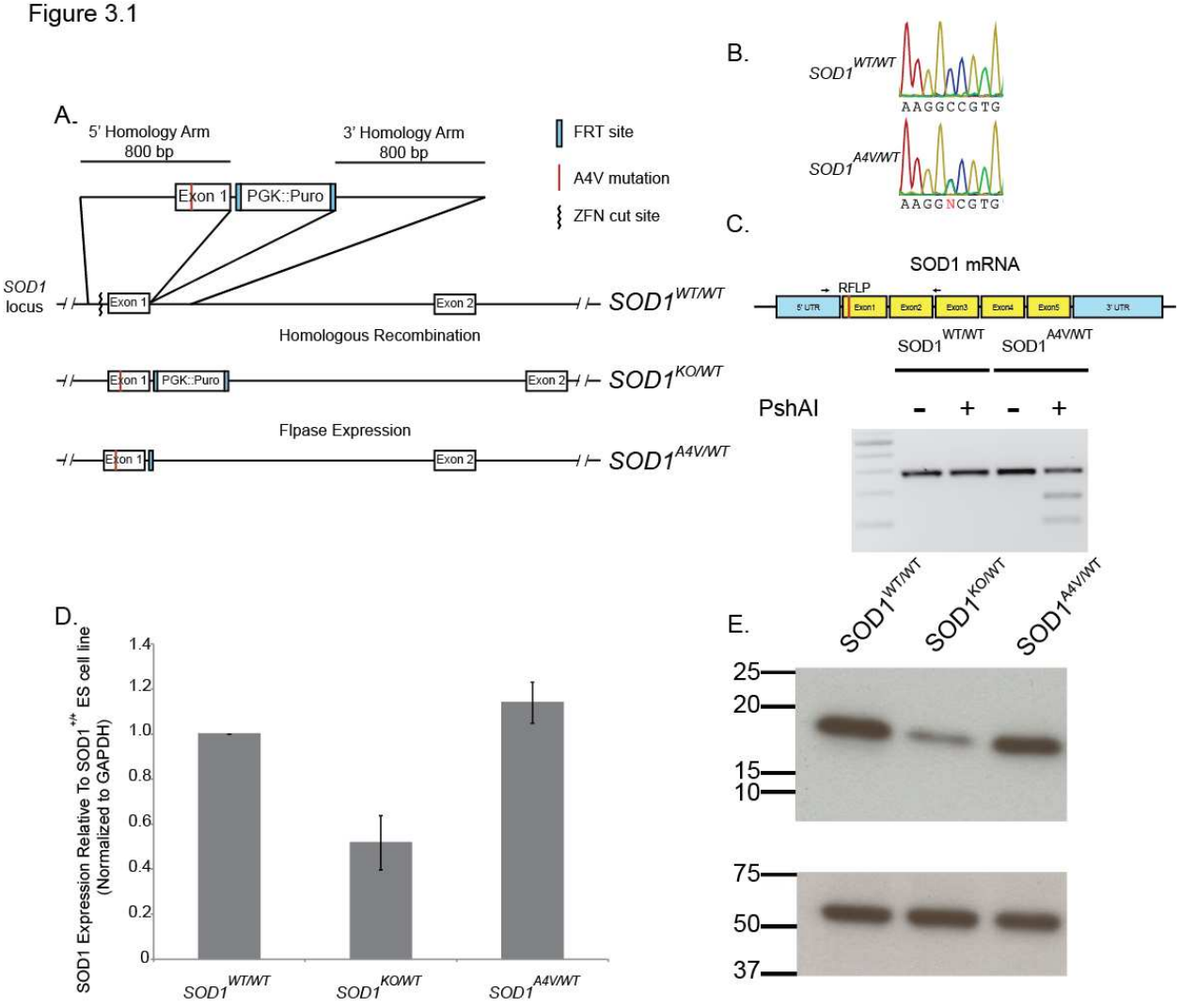

Figure 3.1. The SOD1 A4V mutation was introduced into a control hES cell line. (A) Diagram of gene targeting strategy. (B) Sequencing of SOD1 locus and RFLP analysis for a unique PshAI restriction site (C) confirmed correct targeting of SOD1. (D) qPCR of the gene targeted stem cell lines demonstrated reduced levels of the SOD1 transcript and protein (E) in the intermediate Hues3 SOD1KO/WT ZFN_1 cell line which are restored in the Hues3 SOD1A4V/WT ZFN_1 cell line.

While the fidelity of the ZFNs used in this experiment have been previously validated by whole genome sequencing[9], we decided to generate additional stem cell lines which could be used to demonstrate the sufficiency of the A4V mutation and alleviate the concern of off target mutations introducing modifying mutations elsewhere in the genome. To derive these important controls cell lines we took two strategies. First, using a very similar gene targeting strategy as described previously we corrected the A4V mutation from Hues3 *SOD1*$^{A4V/WT}$ ZFN_1, effectively

reverting the cell line to the original *SOD1 WT* genotype and establishing a second control cell line, from here on referred to as Hues3 $SOD1^{WT/WT}$ ZFNcorr_1 which can be used to determine the impact of any off target mutations introduced during the gene targeting process by comparing Hues3 $SOD1^{WT/WT}$ ZFNcorr_1 to the original untargeted Hues3 $SOD1^{WT/WT}$ cell line (Figure 3.2A). This time, for the editing of the $SOD1^{A4V}$ allele we used the Dre/ROX recombination system to remove the puromycin resistant cassette rather than the FLPo/FRT system used previously due to difficulties in re targeting *SOD1* when the FRT site is already integrated at the genomic locus (data not shown).

Our second strategy was to generate an additional cell line in which the *SOD1 A4V* mutation was introduced using a different site specific nuclease technology. Recently, the use of the bacterially derived CRISPR/Cas9 system has been employed to generate site specific DSBs[15,16]. This nuclease technology is directed to the chosen genomic loci by co-introducing a guide RNA (gRNA) construct which is loaded into the Cas9 protein[16]. The Cas9/gRNA riboprotein complex is then directed to the genomic loci which is complementary to a 20 base pair region of the gRNA[16]. The Cas9 protein then introduces a DSB through its endogenous nuclease activity[16]. By altering the 20 bp sequence of the gRNA which directs the Cas9 protein to genomic DNA it is possible to target virtually any loci within the genome[16,17]. Furthermore, by using a semi-catalytically inactive Cas9 protein it is possible to introduce a single stranded nick at a chosen genomic loci[17]. The use of the Cas9 nickase allows for efficient gene targeting while also limiting the possibility of introducing DSB at off target sites within the genome[17]. Using ZFN and Cas9/CRISPR nuclease technologies to generate separate clones harboring the same introduced mutation should reduce the possibility of common off target mutations driving any phenotypes. We designed gRNAs that would allow for the introduction of a DSB near the 5' of the *SOD1* locus using CRISPR/Cas9 technology. The same gene targeting strategy described above was used to induce the A4V mutation with CRISPR/Cas9 technology except the Cas9 nickase was introduced in the cells to stimulate HR rather than the previously used

ZFN. Several puromycin resistant colonies were established using the Cas9 nickase technology and one was chosen for puromycin cassette removal. After isolation of colonies no longer containing the puromycin cassette a single cell line was chosen for further experiments, here on referred to as Hues3 *SOD1*$^{A4V/WT}$ Cas9n_1. PCR followed by sequencing and RFLP analysis were used to confirm correct targeting of the *SOD1* locus (Figure 3.2B and C). The transcript levels of *SOD1* in the gene targeted cell lines were queried by qPCR to demonstrate expression from the targeted *SOD1* locus (Figure 3.2D).

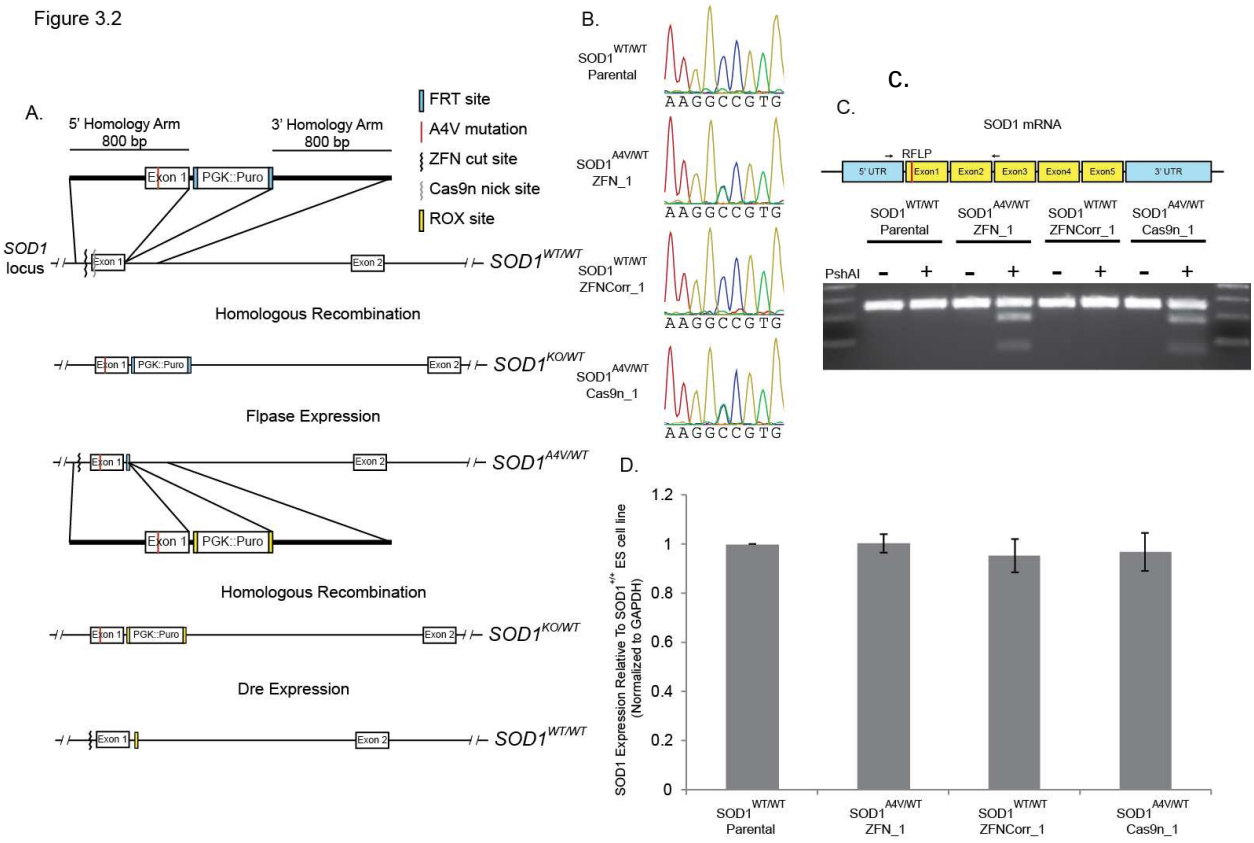

Figure 3.2. Gene targeting of additional cell lines for isogenic controls. (A) Diagram of the targeting strategy. Sequencing (B) and RFLP (C) analysis was used to confirm correct targeting of the SOD1 locus. The SOD1 transcript is expressed at similar levels amongst the gene targeted cell lines when analyzed by qPCR (D).

Currently we are using whole genome sequencing to identify potential off-target mutations and determine the stability of the genomes in the parental Hues3 *SOD1*$^{WT/WT}$, Hues3 *SOD1*$^{KO/WT}$ ZFN_1, Hues3 *SOD1*$^{A4V/WT}$ ZFN_1, Hues3 *SOD1*$^{A4V/WT}$ ZFN_2, Hues3 *SOD1*$^{WT/WT}$ ZFNcorr_1, and Hues3 *SOD1*$^{A4V/WT}$ Cas9n_1 cell lines (Figure 3.3).

Figure 3.3

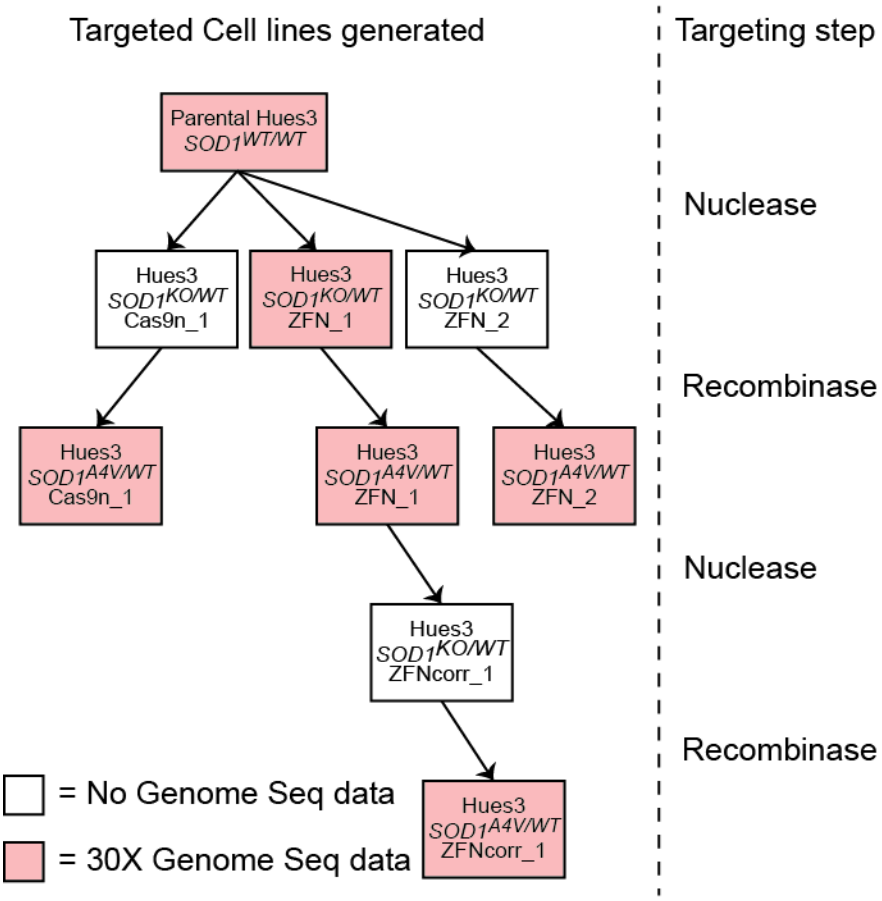

Figure 3.3. Outline of parental and gene targeted hES cells which have been sequenced to over 30X coverage. To the left of the dotted line indicates the cells lines which have been sequenced. To the right of the dotted line indicates the targeting steps which were taken to generate the cell lines. Nuclease = step whereby ZFN or Cas9n was used to modify A4V allele, Recombinase = step whereby FLPo or Dre was used to excise Puromycin cassette from SOD1 locus. Each step required clonal isolation.

Purification of Functional MNs from Isogenic hES Cell Lines. To generate MNs from the parental Hues3 *SOD1*$^{WT/WT}$ and Hues3 *SOD1*$^{A4V/WT}$ ZFN_1 gene targeted stem cell lines, we subjected them to a previously described[9,18] 24 day EB MN differentiation protocol based on six days of dual SMAD inhibition to convert the stem cells to neural progenitors[19] and the addition of RA and SAG for MN specification (Figure 3.4A). On day 24, the cultures were dissociated and FACS was used to analyze the cultures for GFP fluorescence. As described previously[14], when compared to cultures where RA and SAG were omitted from the differentiation protocol, we found a population of cells which exhibited a clear increase in GFP fluorescence (Figure 3.4B).

We used qPCR and Western blotting to confirm the expression and translation of *SOD1* in the differentiated cultures (Figure 3.4C and D). As shown previously[14], the cells which expressed GFP were found to also express Islet1 when queried by immunostaining (Figure 3.4E). We found it was possible to plate the MNs down on poly-d-Lysine and Laminin coated tissue culture plates after purification, allowing for the establishment of cultures composed entirely of MNs, without the influences of other neuronal cell types (Figure 3.4F).

Figure 3.4

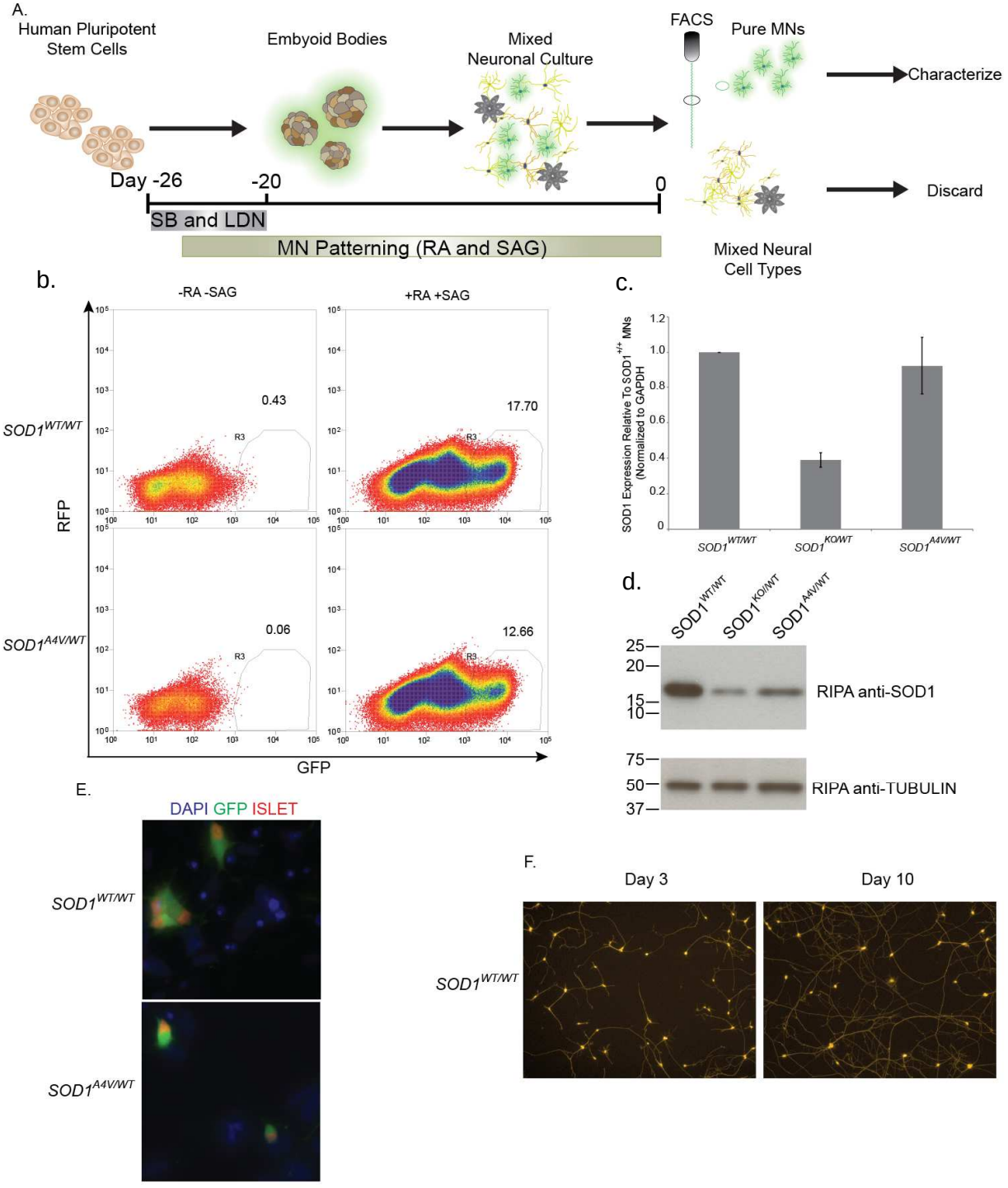

Figure 3.4. Parental and gene targetd hES cell lines can be differentiated into MNs and purified by FACS. (A) Schematic of EB differenitation protocol with FACS isolation of MNs. There is a clear population of cells which express GFP as determined by FACS (B). The neurons generated from this protocol express the SOD1 transcript (C) and protein (D). The GFP-positive cells co stain for ISLET indicating MN identity (E). The MNs can be plated after FACS isolation and maintained as purified MN cultures (F) as determined by live imaging of MN cultures stained with a live dye at sequential time points.

SOD1 Aggregation. Aggregation of SOD1 is a common pathological finding in ALS patients with familial *SOD1* mutations[2]. While the clinical significance of these aggregates is controversial, they are nonetheless an important demonstration of the altered processing of SOD1 mutant proteins[2]. Our previous findings demonstrated large scale aggregates were not necessary for the degeneration of MNs in our culture system however, insoluble SOD1 could be observed after treatment with the proteasome inhibitor MG132[9]. To determine the sufficiency of the *SOD1 A4V* mutation to alter the solubility of SOD1 within in vitro derived MNs, we generated MNs from the Hues3 $SOD1^{WT/WT}$ and Hues3 $SOD1^{A4V/WT}$ ZFN_1 cell lines using the 24 day EB protocol. Because it was difficult to generate and FACS purify the required number of MNs needed for biochemical studies, the unpurified differentiated cultures were plated after dissociation and allowed to mature for 10 days in vitro. On day 10, the cultures were treated with the proteasome inhibitor, MG132, or DMSO as a control. After 5 days of treatment, soluble protein was extracted from the cultures using RIPA lysis and solubilization buffer (Figure 3.5A). The cell lysate was then centrifuged and the supernatant was transferred into a separate tube. The remaining insoluble protein was washed with PBS, then centrifuged to pellet the protein. Urea buffer was then used to solubilize the remaining protein pellet. When western blotting was used to probe for SOD1 in the soluble and insoluble protein fractions, we found insoluble SOD1 was only observed in cultures which were both expressing the SOD1 A4V protein and treated with the proteasome inhibitor (Figure 3.5B). While we cannot rule out low levels of insoluble SOD1 which could not be detected using our methods, this finding suggests that large-scale SOD1 aggregates do not develop in the Hues3 stem cell-derived MNs under normal conditions, even when expressing a mutant form of the protein. This result demonstrated the sufficiency of the *SOD1 A4V* mutation to alter the processing of the SOD1 protein in a similar manner to our previous findings using MNs derived from patient specific iPSCs. Because of the concordance of our results from the previous study using patient specific iPSCs[9] and the current study using gene targeted hES cell lines we pursued our investigation of other phenotypes we had

previously demonstrated.

Figure 3.5

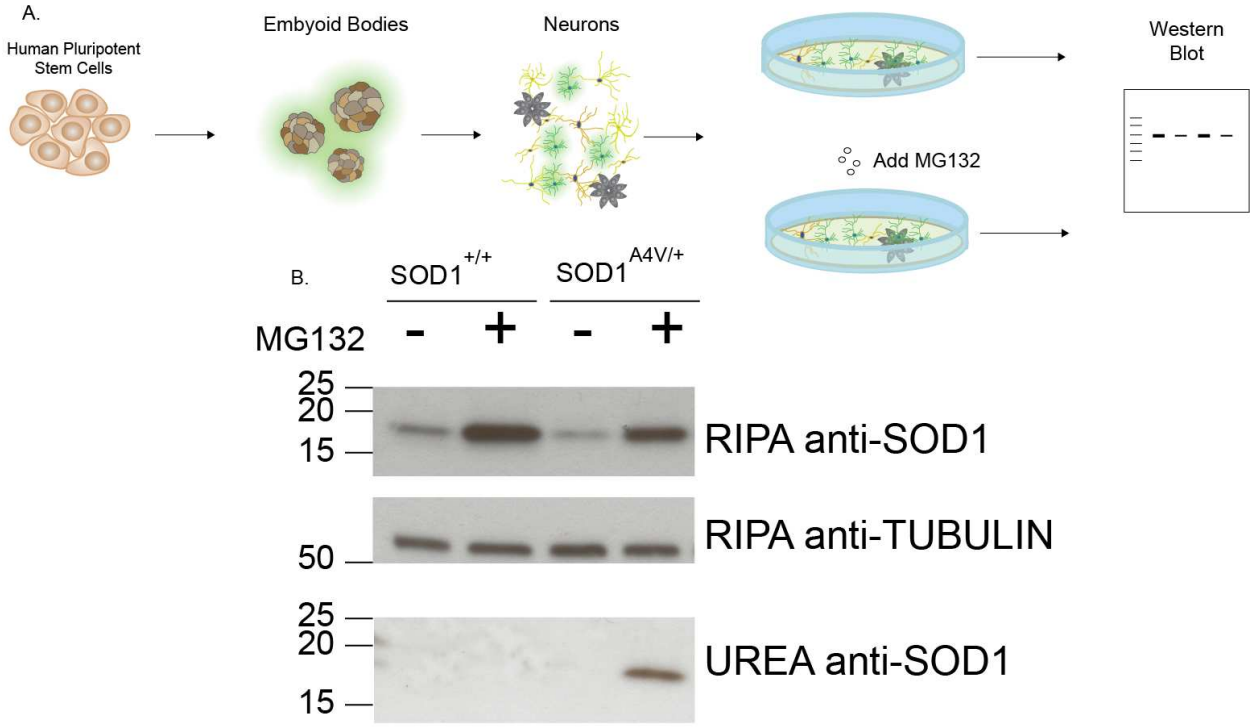

Figure 3.5. Solubility status of SOD1WT/WT and SOD1A4V/WT in neuronal cultures. (A) Diagram of experiment. SOD1 accumulates in the insolube protein fraction (B) when neuronal cultures are expressing the SOD1 A4V variant and are treated with the proteosome inhibitor MG132.

Survival of MNs expressing the *SOD1 A4V* variant. We previously found that MNs derived from patient specific iPSCs harboring the *SOD1 A4V* mutation exhibited higher levels of apoptosis than control MNs[9]. Importantly, the reduced survival was rescued upon correction of the A4V mutation demonstrating the necessity of the mutation for the diminished survival[9]. We were interested in determining the impact of the induced *SOD1 A4V* mutation on the survival of MNs derived from the gene targeted hES cell line. We used the parental Hues3 $SOD1^{WT/WT}$ and HuES3 $SOD1^{A4V/WT}$ ZFN_1 cell lines to generate MNs which were then purified by FACS (Figure 3.6A). Removing all non-MNs expressing the *SOD1 A4V* protein from the cultures should reduce the chances of non-cell autonomous toxicity impacting MN survival. The isolated MNs were plated on a monolayer of primary glia and the cultures were fixed and stained for human

nuclei and TUJ1 on days 3 and 30 (Figure 3.6B). Whole well imaging and automated counting was used to determine the number of MNs remaining after 3 and 30 days of culture. A trend for reduced survival was found in MNs expressing the *SOD1 A4V* protein when compared to the survival of MNs derived from the control cell line (Figure 3.6C n=4, p=0.09).

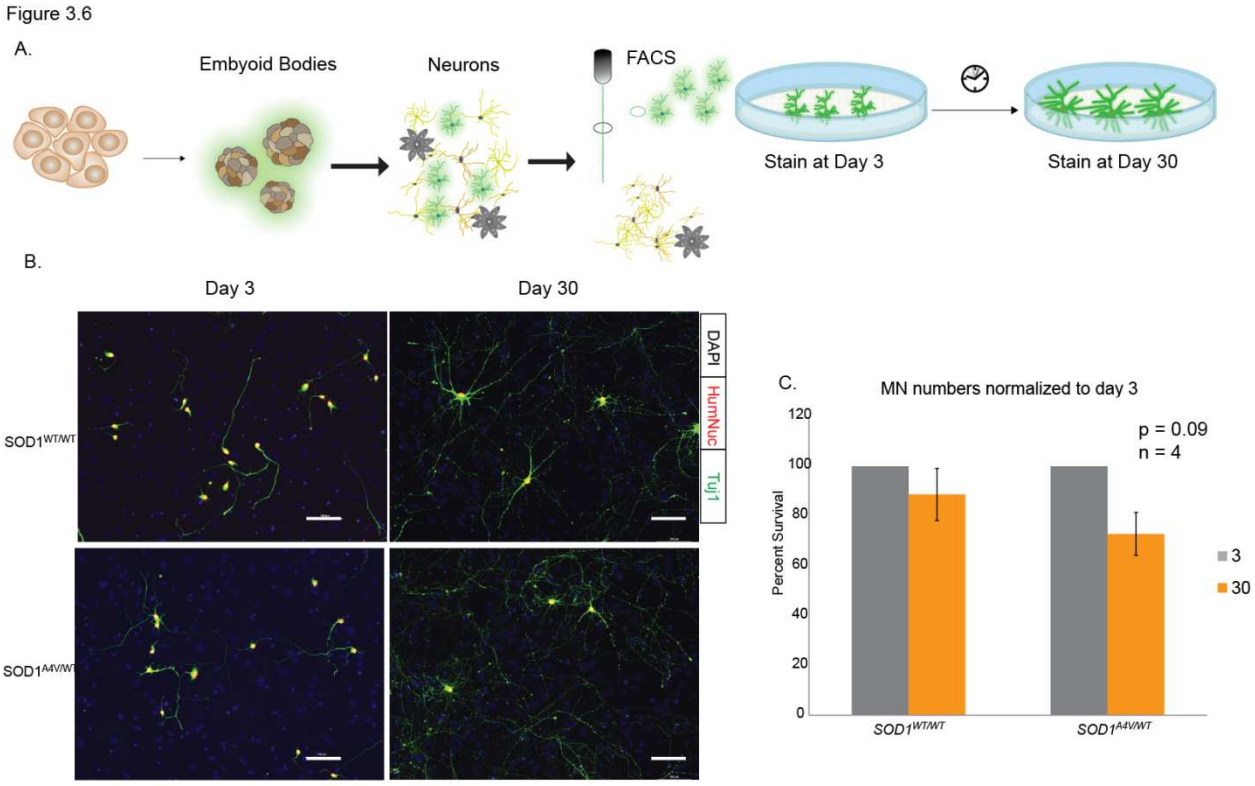

Figure 3.6. There is a trend for reduced survival in MNs expressing the SOD1 A4V variant when compared to controls. (A) Schematic of experiment. (B) Representative images of Human Nuclei and TUJ1 staining at days 3 and 30. (C) A non-significant reduction in MN survival is found in the SOD1A4V/WT MNs when compared to WT controls. (p = 0.09, n= 4)

It is currently unclear why the reduced survival of MNs expressing the *SOD1 A4V* allele was not as prominent in this experiment when compared with our previous findings. It is possible the initial FACS purification of the MNs sub selects for particularly hardy MNs, as they must survive the stressful experience of dissociation and subsequent exposure to high pressure in the FACS machine. Additionally, the use of FACS purified MNs eliminated other cell types which could be expressing the *SOD1 A4V* protein. In the original MN survival experiments[9], other non-MN cell types expressing the *SOD1 A4V* protein may have induced non cell

autonomous toxicity in the MNs, resulting in the observed survival deficit[9]. Another explanation could be the presence of modifying mutations in the Hues3 genetic background, which confer partial resistance to the apoptotic effects of the *SOD1 A4V* variant[2]. To address these possibilities, we are currently performing the MN survival assay described in Chapter 2[9] with the Hues3 isogenic cell lines, in which the MNs are not FACS purified but instead the entire differentiated mixture is plated on murine glia and MNs are identified and counted by Tuj1 and Islet co-staining.

RNA Sequencing of EB Derived and Purified MNs Expressing WT or A4V *SOD1*. The previous study using patient-specific iPSCs used transcriptional profiling by RNA sequencing to uncover pathways modulated by expression of the *SOD1 A4V* protein[9]. These experiments proved enlightening, but were performed on MNs which had been cultured in the presence of other cell types which expressed the *SOD1 A4V* protein, leaving open the possibility that non-MN cell types in the culture could have induced transcriptional changes in a non-cell autonomous manner[14]. To eliminate the confounding influence of other cell types expressing the mutant protein in differentiated cultures, we used the parental Hues3 $SOD1^{WT/WT}$ and Hues3 $SOD1^{A4V/WT}$ ZFN_1 hES cell lines to generate MNs which were then isolated by FACS and plated as pure MN cultures on poly-d-lysine and laminin. After 15 days of in vitro culture with MN media which had been conditioned in the presence of murine glial cells, RNA was extracted from the MN cultures and used to generate libraries which were subjected to next generation sequencing (Figure 3.7A). The sequenced reads were aligned to the human genome, and differentially expressed genes were identified by the tuxedo suite of bioinformatics tools. Unsupervised hierarchical clustering segregated the samples based on expression of the SOD1 WT or A4V proteins, demonstrating that the mutant protein was capable of driving transcriptional changes in our in vitro derived MNs (Figure 3.7B).

Figure 3.7

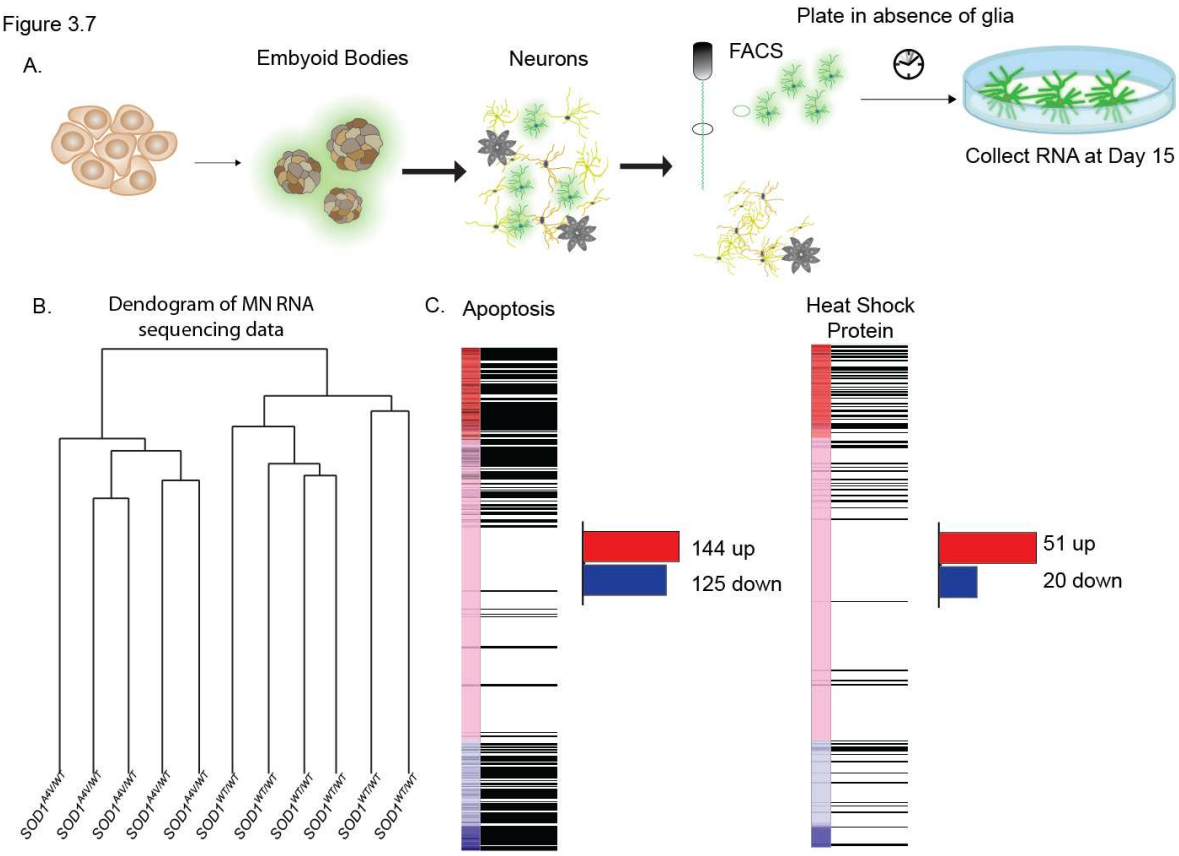

Figure 3.7. Transcriptional profiling of purified parental Hues3 SOD1WT/WT and Hues3 SOD1A4V/WT MNs. (A) Diagram of experiment. Unsupervised hierarchical clustering separates the samples based on genotypes (B). Gene Set Analysis Enrichment identifies the Apoptosis pathway and Heat Shock Proteins as being significantly upregulated (C) (FDR q-value < 0.01).

A major question we wanted to answer was how similar the transcriptional differences found in the current study would be to our previous RNA sequencing experiments[9]. We found the transcriptional differences found in the current study were not in concordance with the altered transcripts found in our previous transcriptional profiling experiments using MNs derived from the patient-specific iPSC and its gene corrected control. While there was no clear overlap between the two data sets in regards to gene by gene comparisons, there were signals of degenerative processes being underway in the current data set when analyzed by Gene Set Enrichment Analysis (GSEA)[20]. For example, genes involved in apoptosis and the heat shock response were found to be significantly up regulated (Figure 3.7C FDR q-value < 0.01). As stated in regards to the MN survival data, in which there was a nonsignificant trend for reduced

survival of MNs generated from the Hues3 *SOD1*$^{A4V/WT}$ ZFN_1 hES cell line when compared to the parental Hues3 HB9:GFP cell line, there are several reasons this may have occurred. First, the different genetic backgrounds may influence the impact of the *SOD1 A4V* protein in different ways[2]. Second, the current study used glial-conditioned MN media rather than plating in the presence of glial cell types to facilitate the isolation of human MN RNA without murine glial contamination. It is possible that direct contact with glial cell types alters the maturation of synaptic networks of MNs, and results in large differences in the transcriptional profiles of MNs in the presence vs. absence of glial cells. To come to a more uniform understanding of transcriptional changes driven by the *SOD1 A4V* variant, we are currently performing transcriptional profiling on the parental Hues3 *SOD1*$^{WT/WT}$ and Hues3 *SOD1*$^{A4V/WT}$ ZFN_1 as well as the previously described 39b *SOD1*$^{A4V/WT}$ and gene corrected 39b *SOD*$^{1A4V/WT}$ cell lines in the exact manner described in Chapter 2. Additionally, we recently generated a third set of isogenic *SOD1*$^{WT/WT}$ and *SOD1*$^{A4V/WT}$ pairs by correcting the *SOD1 A4V* mutation from the previously described Rb9d iPSC cell lines[9]. These iPSCs will also be used for transcriptional profiling as described in Chapter 2. We hope that by using 3 different sets of cell lines, each of which have undergone gene targeting to generate genetically controlled comparison cell lines, we can identify robust and meaningful gene expression differences. Furthermore, due to technical difficulties in the implementation of the EB protocol for the production of large numbers of MNs, which are necessary for RNA sequencing experiments, we adopted an alternative MN directed differentiation protocol, based on approaches recently described by Lorenz Studer's research group[21]. The bulk production and purification of MNs allows for testing many more conditions than were previously possible, which holds promise for clarifying previous conflicting results.

Bulk Preparation of ES Cell Derived MNs. To facilitate bulk preparations of purified MNs, we used an adherent MN differentiation protocol. The adherent protocol reflects similar developmental pathways as the EB protocol, including dual SMAD inhibition to differentiate the stem cells to neural progenitors[19], along with the addition of RA and SAG to caudalize and

ventralize the neurons into a MN identity. However, after day 5 of differentiation we also included the small molecule DAPT, a notch inhibitor, and SU5402, an FGF inhibitor. This modification was based on a recent study demonstrating an increased efficiency of neuronal differentiation and decreased time to neurogenesis when these small molecules are included in a directed differentiation[21]. The inclusion of DAPT and SU5402 allowed for a dramatic decrease in the time required for MNs to develop (Figure 3.8A). This allowed for robust production of MNs in 14 days, a 10 day decrease in protocol length when compared to the original EB protocol. When FACS purification was used to monitor GFP fluorescence, a clear population of GFP positive cells could be identified when compared to cultures where RA and SAG were omitted from the differentiation protocol (Figure 3.8B). When plated after FACS purification, a majority of the isolated GFP positive cells were found to be immunopositive for ISLET1 (Figure 3.8C and D). A significant difference in differentiation efficiency was not found between the parental and genome edited stem cell lines, indicating the absence of developmental complications driven by the expression of the *SOD1 A4V* variant (data not shown).

Figure 3.8

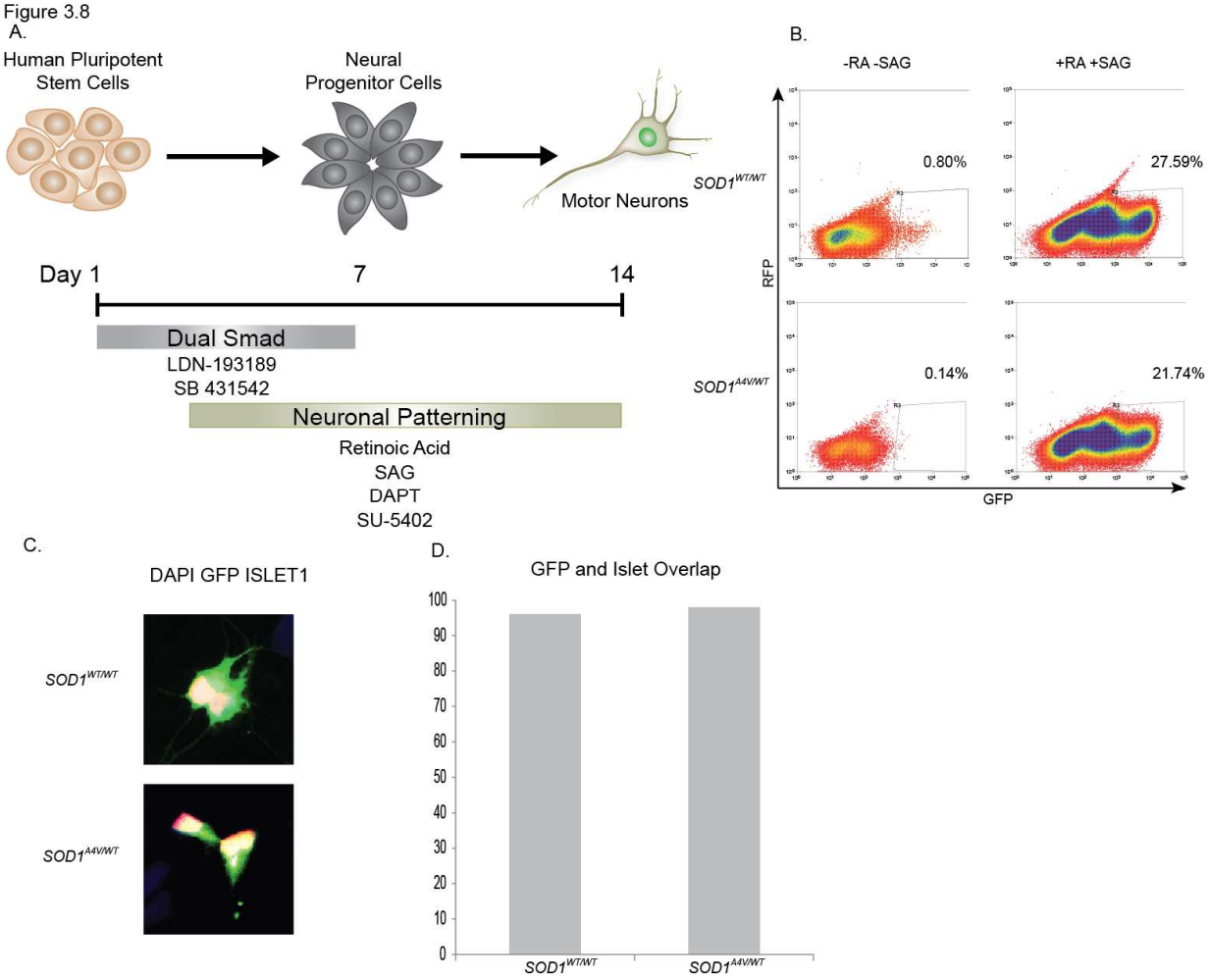

Figure 3.8. Bulk purification of MNs. (A) Protocol for the adherent differentiation of MNs. A clear GFP-positive population of cells is present after 14 days of differenitation as determined by FACS (B). The purified GFP-positive cells co-stain for Islet1 (C and D) indicating MN identity (n = 100 neurons for each genotype, GFP and Islet1 overlap = 96% for the parental Hues3 SOD1WT/WT and 98% for Hues3 SOD1A4V/WT ZFN 1).

Mitochondrial Swelling in Purified MNs Harboring the *SOD1 A4V* Allele. We had previously found malformed mitochondria within cells expressing the *SOD1 A4V* variant when investigated by electron microscopy (EM)[9]. In that study, we found the mitochondrial defect was removed upon genetic correction of the *SOD1* A4V mutation, demonstrating the necessity of the mutation for the phenotype[9]. To demonstrate the sufficiency of the A4V mutation to drive this mitochondrial defect, we used the adherent differentiation protocol to generate MNs from 1) the parental HuES3 *SOD1*$^{WT/WT}$, 2) Hues3 *SOD1*$^{A4V/WT}$ ZFN_1, 3) Hues3 *SOD1*$^{A4V/WT}$ Cas9n_1n, and 4) Hues3 *SOD1*$^{WT/WT}$ ZFNcorr_1, then FACS purified MNs based on GFP fluorescence. Twenty

thousand MNs were plated on coverslips and cultured with glia-conditioned MN media (Figure 3.9A). On days 20 and 30, the MNs were fixed and processed for EM. There were no apparent differences between *SOD1 WT* and *SOD1 A4V* in the shape of mitochondria from MNs cultured for 20 days (data not shown). However, swollen mitochondria were found in MNs which expressed the *SOD1 A4V* variants and had been cultured for 30 days (Figure 3.9B). This finding demonstrates the sufficiency of the A4V mutation for mitochondrial swelling as well as the degenerative nature of this defect, as it was seen in MNs cultured for 30 days but not 20. This finding also importantly depicted mitochondrial swelling within purified MNs in the absence of other cell types which could be influencing the phenotype, which was technically impossible to confirm in our previous experiments.

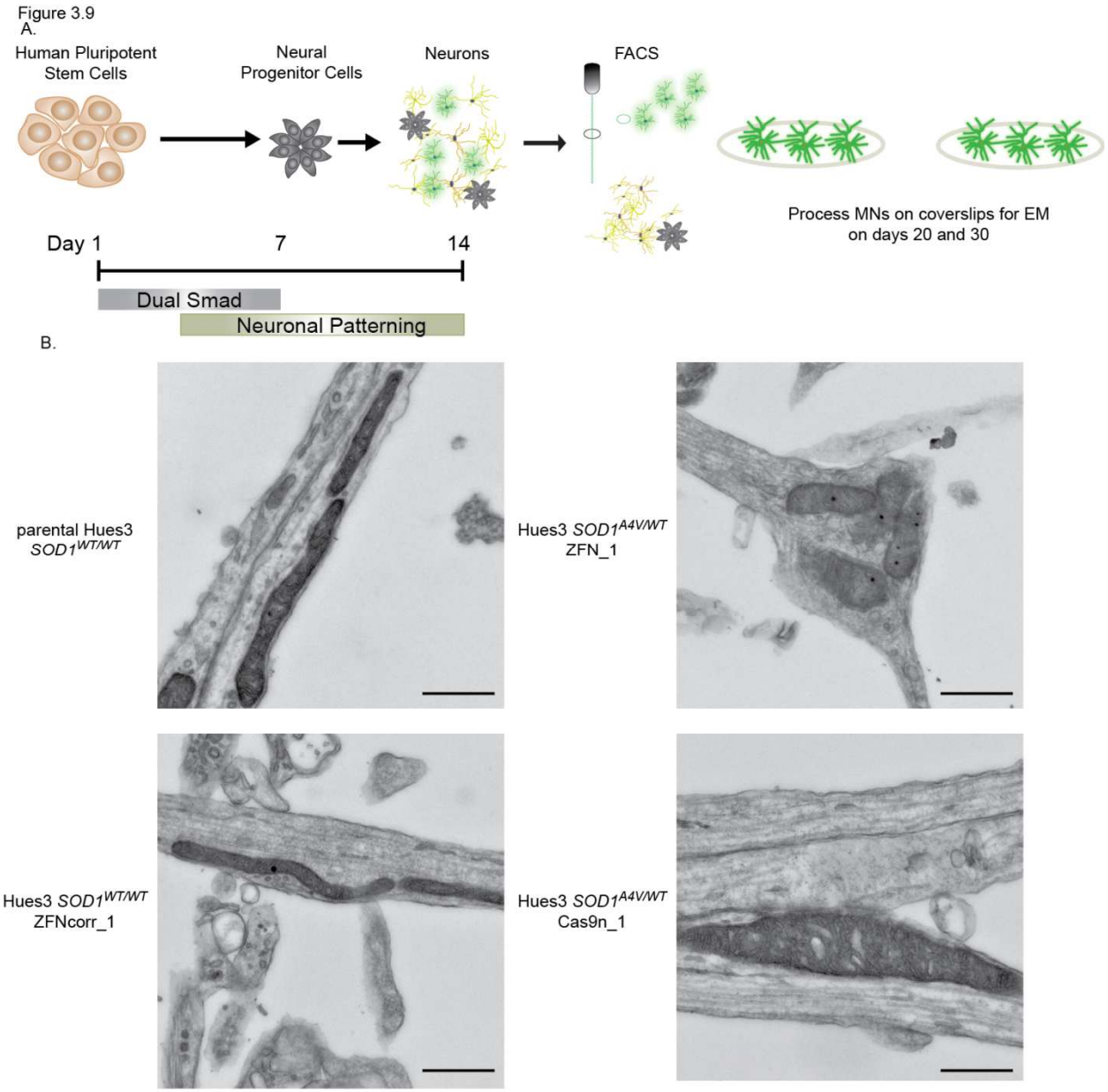

Figure 3.9. Malformed mitochondria are found in MNs derived from gene targeted Hues3 cell lines harboring the SOD1 A4V mutation. (A) Diagram of the experiment. (B) Only MNs which are expressing the A4V variant contain swollen mitochondria (scale bar = 500nm).

Hyperexcitability in Purified MNs Expressing the *SOD1 A4V* Protein. To investigate the excitability of the purified MNs from the parental and gene targeted stem cell lines, the adherent differentiation protocol was used to derive MNs from 1) the parental HuES3 *SOD1*$^{WT/WT}$, 2) Hues3 *SOD1*$^{A4V/WT}$ ZFN_1, and 3) Hues3 *SOD1*$^{A4V/WT}$ Cas9n_1 hES. FACS was used to isolate GFP+ cells, which were then plated on 12-well MEA plates with murine glia. The

purified MNs were allowed to mature in culture while MEA recordings were taken every 3 days (Figure 3.10A). The cultures of MNs expressing the SOD1 A4V protein exhibited heightened spike rates after three weeks of culture, mirroring our previous findings with unsorted MNs[9,22](Figure 3.10B). This result supports our previous interpretation that the increased excitability in MNs expressing the mutant SOD1 protein was not derived from synaptic inputs of other inhibitory or excitatory neurons on the MNs[22], as the MNs in the current experiment were purified and plated in the absence of other neuronal subtypes. Furthermore, the current study demonstrates the sufficiency of the induced A4V mutation to drive the heightened excitability in purified MNs. Lastly, the absence of other cell types expressing the SOD1 A4V protein indicates that the increased firing rate results from the impact of the mutant protein in the MNs themselves, rather than from the influence of non-MN cell types expressing the A4V variant.

Figure 3.10

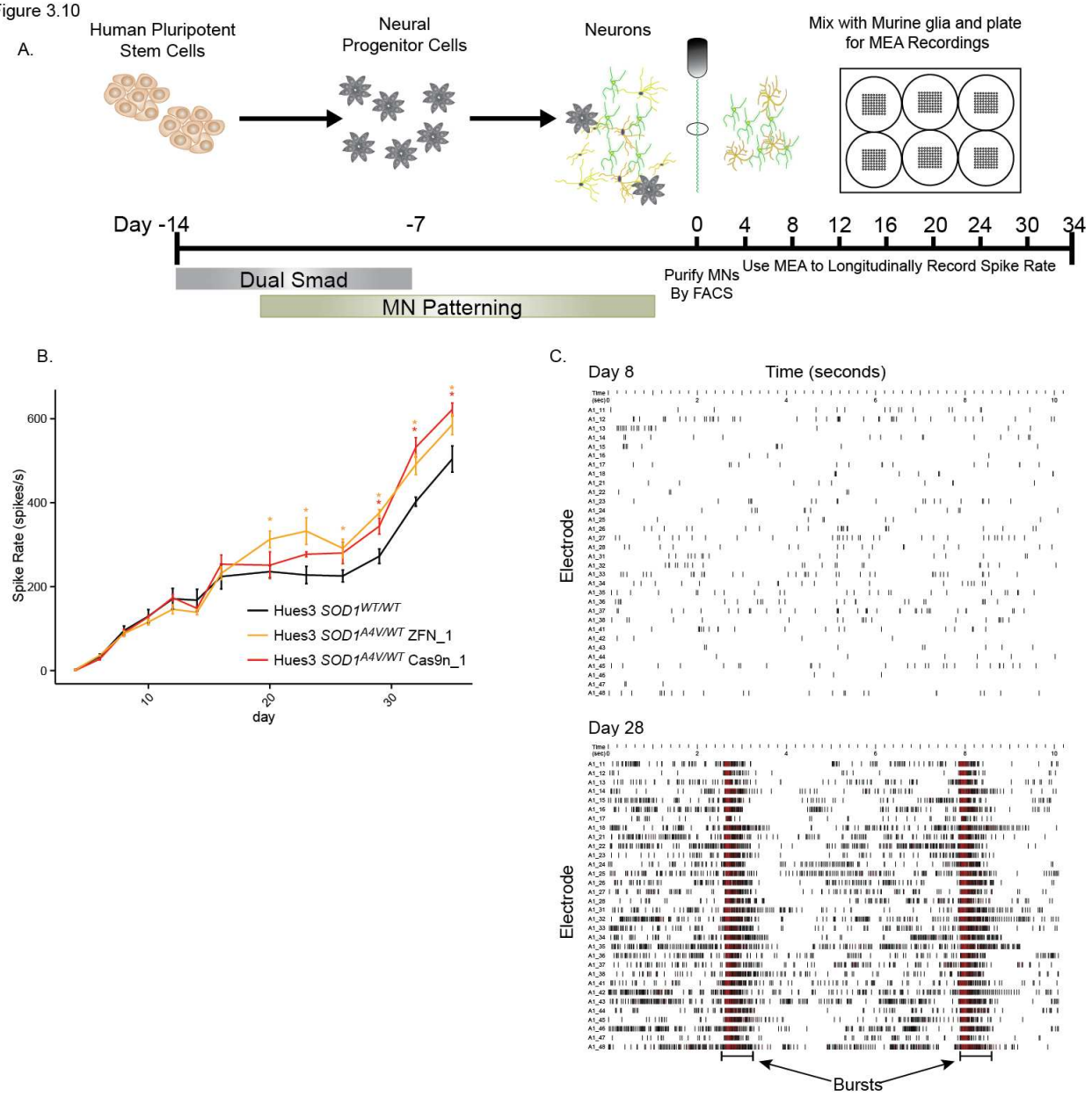

Figure 3.10. MNs derived from the gene targeted Hues3 cell lines which express the SOD1 A4V transcript show increased excitability. (A) Diagram of experiment. (B) Longitudinal MEA recordings demonstrate hyper-excitability of Hues3 SOD1A4V/WT ZFN_1 and Hues3 SOD1A4V/WT Cas9n_1 derived MNs ($p < 0.05$). (C) Raster plots demonstrating synchronized "bursting" activity in mature cultures.

When investigating neuron excitability using the MEA, we observed that the neurons began to fire synchronously after approximately 10 days in culture (Figure 3.10C). Though this "bursting" activity was common to all genotypes we are currently analyzing the impact of the SOD1 A4V protein on this network activity

RNA Sequencing of Purified MNs from Sequential Time Points. Longitudinal MEA recordings demonstrated dramatic changes in physiology as the MNs matured in culture. We were interested in acquiring a global perspective of the dynamics of MN maturation in vitro, and how this process may impact disease-relevant transcriptional signatures. To this end, we used RNA sequencing after multiple sequential durations of culture to develop a time course of in vitro MN maturation and *SOD1 A4V* genotypic deviation. To acquire the large number of MNs needed for harvesting RNA at multiple time points, we used the adherent differentiation protocol described above to generate MNs from the parental Hues3 $SOD1^{WT/WT}$ and Hues3 $SOD1^{A4V/WT}$ ZFN_1 lines, which were purified using FACS based on GFP expression. The purified MNs were plated on poly-d-lysine/laminin coated plates and cultured with glia conditioned MN media. Every other day from day 2 onward, RNA was extracted by adding trizol directly to the culture wells, thus preserving any axonally transported mRNA which may be eliminated if neuronal cultures were dissociated before the isolation of RNA (Figure 3.11A). Sequencing libraries were prepared from the isolated RNA using the TruSeq sequencing library preparation kit and sequenced on an Illumina HiSeq 2500. Tuxedo suite bioinformatics tools were used to align the sequencing reads and identify significantly differentially expressed transcripts.

Unsupervised hierarchical clustering revealed grouping based on chronological age of the MNs as well as the *SOD1* genotype (Figure 3.11B). This confirmed our assumption that the amount of time spent in culture has a major impact on the gene expression profile of the MNs. The time point exhibiting the largest number of genotypically altered transcripts was day 4 (Figure 3.11C). Furthermore, the time point at which the most unique gene expression changes, meaning these genes were not differentially regulated between $SOD1^{WT/WT}$ and $SOD1^{A4V/WT}$ MNs at any other time point was day 4 followed by day 2. This finding was at first surprising, as we had predicted that genotypic differences would increase over time due to the degenerative nature of the *SOD1 A4V* mutant protein. However, it is possible that the increase in the genotypic differences at the early time points is partially driven by the dynamic nature of gene

expression found during this period. In support of this hypothesis, we also found that when investigating sequential time points for differentially expressed transcripts, days 2 to 4 had the highest number of altered expression levels, followed by days 4 to 6 (Figure 3.11D). This finding highlights the importance of carefully considering the time points for transcriptional investigation, as considerable variability can be introduced from stochastic differences in neuronal maturation, which may be independent of genotype. Furthermore, more biological replicates will be necessary to confirm any genotypically differentially expressed transcripts, if they are found at earlier and thus potentially more variable time points.

Figure 3.11

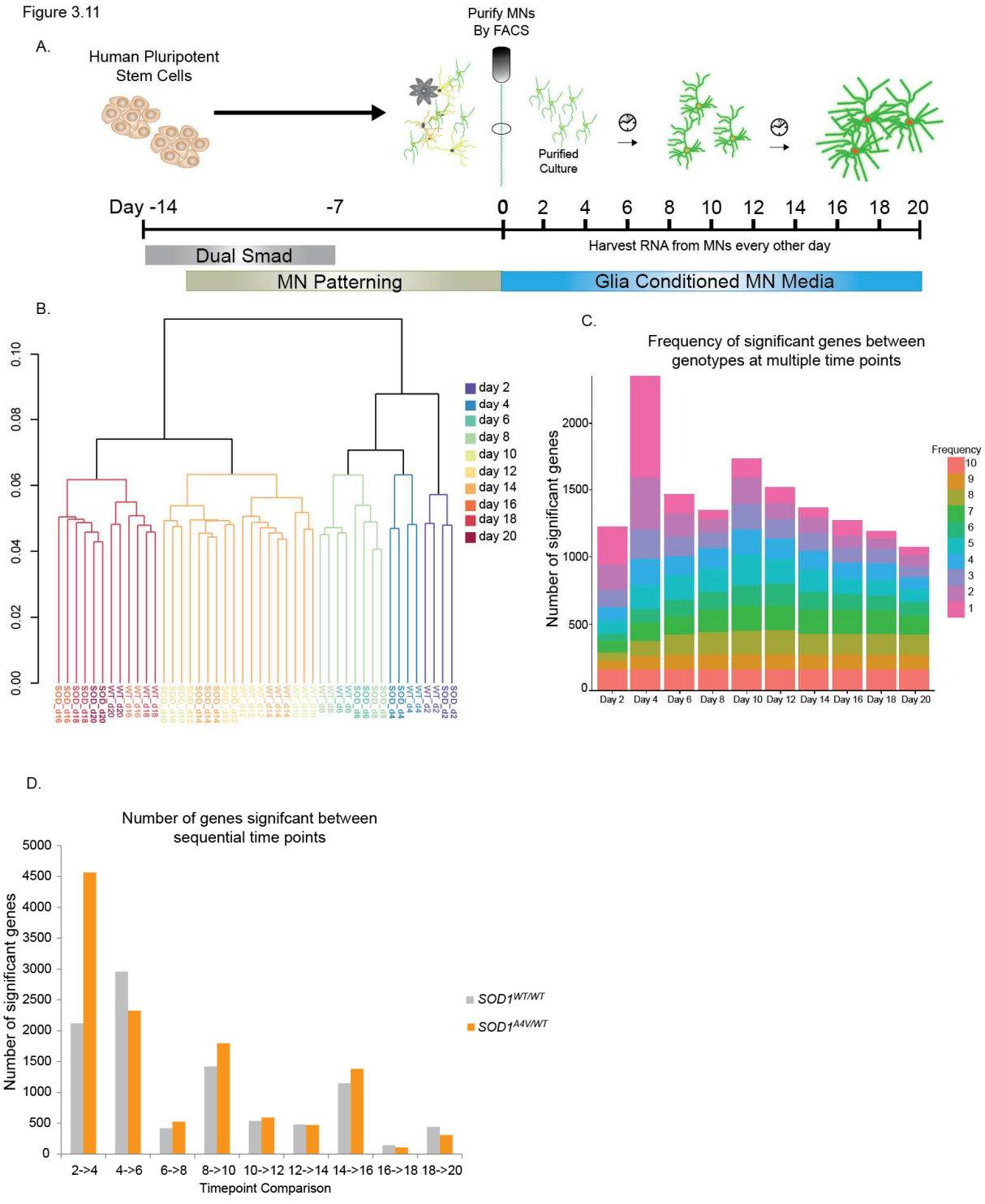

Figure 3.11. RNA sequencing of MNs at sequencial time points. (A) Diagram of experiment. (B) Unsupervised hierarchical clustering segregates the MNs by chronological time and genotype. (C) The time point which contained the most unique genotypically altered transcripts was day four and the time point with the least was day 20. (D) The number of significantly altered transcripts between sequential timepoints in SOD1WT/WT MNs is from day four to six and in SOD1A4V/WT MNs is from day two to four indicating the initial days in culture are transcriptionally dynamic.

When investigating sets of genes with common cellular function, genes involved in synaptic transmission were found to be upregulated over time in both genotypes, likely representing the maturation of the neurons in our cultures. Specifically, transcripts of the synaptotagmin and voltage gated ion channel families were significantly and continuously increased throughout the time course, indicating the establishment of synapses and development of electrophysiological capacity in MNs within the cultures (Figure 3.12A FDR q-value < 0.01). Previously, we had found gene sets associated with ER stress and mitochondria to be misregulated in MNs expressing the SOD1 A4V protein[9]. In the currently described data set, these pathways were not found when GSEA weas used to query the RNA sequencing information. While widespread increases in ER stress were not found, when we investigated the ratio of spliced to unspliced XBP1, an initial step in the initiation of the unfolded protein response[23], there was a non-significant trend for an increase in this ratio on day 20 when MNs expressing the A4V protein were compared to controls (Figure 3.12B).

Figure 3.12

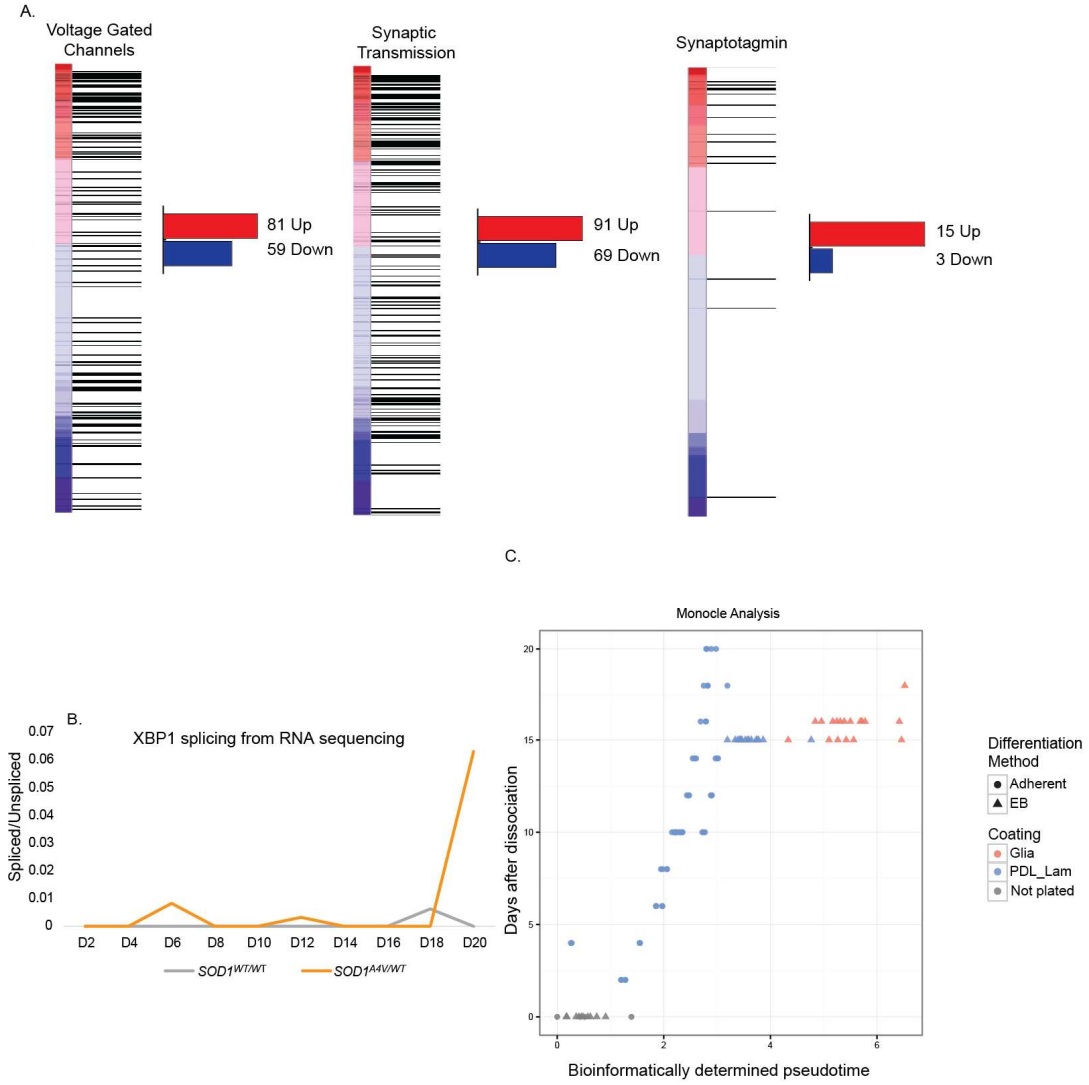

Figure 3.12. Analysis of sequential RNA sequencing experiment reveals maturation of MN cultures. (A) Gene set enrichment analysis found pathways involved in electrophysiological and synaptic maturity to be up regulated in Hues3 SOD1WT/WT MNs between days 2 and 20 (FDR q-value < 0.01) (B) A trend for increased XBP1 splicing in Hues3 SOD1A4V/WT ZFN_1 derived MNs when compared to parental Hues3 SOD1WT/WT MNs was found only at day 20. (C) Bioinformatic analysis of multiple RNA sequencing data sets demonstrates the influence of culture conditions and differentiation method on the maturation of MNs in vitro.

Several factors can explain why the same gene expression signatures were not found when RNA sequencing was used to identify differentially expressed genes in MNs derived from *SOD1 WT* and *SOD1 A4V* pluripotent cell lines. First, in the previous experiments using patient derived iPSCs, the entire differentiated culture was plated on glia and cultured for 15 days before FACS purification of the MNs. This allows for the possibility of many other cell types to

interact with the MNs during their time in culture, which may facilitate increased maturation of the MNs or cause more rapid degeneration due to non-cell autonomous toxicity of other cells expressing the SOD1 A4V protein in culture[12]. Furthermore, the EB protocol was used to generate MNs for the previously described RNA sequencing experiment, which is 10 days longer and thus may produce more mature MNs which exhibit signs of degeneration earlier than MNs produced from the adherent protocol. To test these possibilities, we used the bioinformatics tool MONOCLE to compare several MN RNA sequencing data sets generated by different members of our group. Results indicated that the MNs used in the currently described RNA sequencing experiment, which were purified at day 0 and cultured in the absence of all other cell types, were bioinformatically less mature than the MNs used in the previously described RNA sequencing experiments which were purified after being cultured with murine glia for 15 days[9] (Figure 3.12C). Additionally, transcriptional signatures derived from expression of the *SOD1 A4V* variant may be confounded by differences in genetic modifiers between the cell lines used[1]. Currently, we are producing RNA sequencing libraries from MNs which were cultured for longer durations of time than used in this study in the hopes that more mature degenerative signatures can be revealed.

While the current RNA sequencing data are not entirely consistent with our previous data, the results were informative, and served to shift our focus for next step experiments. We decided to focus on the genes which were most likely to alter the physiological properties of the neurons. To operationally define these genes, we first used 5 FPKM as a cutoff for expression, based on an earlier finding that genes expressed below these levels are often difficult to detect and therefore reproduce (unpublished finding). Transcripts which had previously been shown to encode delayed rectifying potassium channels were prioritized even if not found to be differentially expressed, as we previously found deficits in delayed rectifying potassium currents in *SOD1*$^{A4V/WT}$ MNs[22]. Because the dynamic nature of neuronal maturation in the initial time points can lead to irreproducible expression differences, we eliminated genes which were not

significantly altered between the *SOD1 WT* and *SOD1 A4V* MNs at the day 20 time point. Considering that the MNs did not show increased excitability until day 20, we believe that the differences directly preceding this event are likely the most important. Other transcripts which did not encode voltage-gated channels but were differentially expressed between the two genotypes and had some implication in excitability or neurodegeneration were also chosen as interesting targets (Figure 3.13A). We performed qPCR on multiple biological replicates generated in the exact same manner as the material used for RNA sequencing, to determine the reproducibility of gene expression differences in the initial list of chosen targets (Figure 3.13B). We used the qPCR results to further eliminate genes which did not produce robust alteration in gene expression when compared to the RNA sequencing data. We confirmed the differential expression in 35% of the targets (9 of 26).

Figure 3.13

A.

Target genes from RNA sequencing data

KCNV1	*NBL1*	*KCNQ3*	*DNAJA4*
KCNH2	*PSMG3*	*LY6E*	*EDNRB*
HCN1	*ABCA1*	*LY6H*	*MC4R*
KCNQ2	*IGFBP4*	*TCEAL5*	*GPR37*
PCP4	*GRM3*	*TCEAL6*	*PTGER3*
SEZ6	*PTPRT*	*KCNJ2*	*SLC44A5*
KCNIP1	*CHCHD2*	*FGF13*	*SLC24A3*
SV2A	*SLC10A4*	*SLC8A3*	*ARIH2*

B.

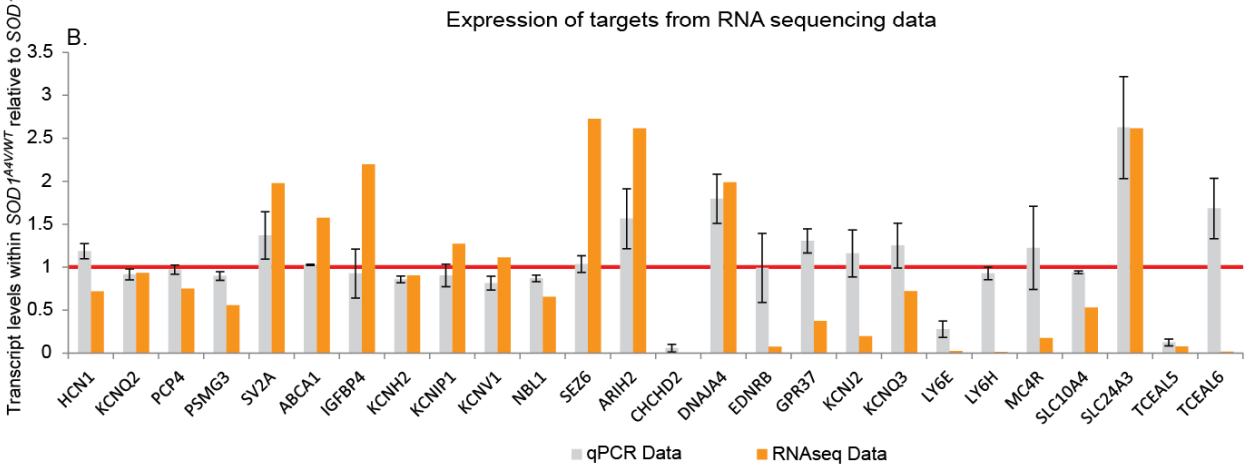

Figure 3.13. (Continued) A group of target genes was identified which could be modulators of phenotypic differences between parental Hues3 SOD1WT/WT and Hues3 SOD1A4V/WT derived MNs. (A) A list of genes was generated based on differential gene expression or being implicated in physiological activity. (B) qPCR was used to confirm differential expression of the target genes (n = 2, genes which we could not obtain qPCR data are not shown).

Electrophysiological Measurements after Knockdown of Target Genes in Stem Cell-Derived MNs. As the next step in our analysis, we investigated the impact of reducing the levels of several genes found to be expressed in the stem cell-derived MNs from our RNA sequencing data. To this end, we planned to use an siRNA approach to down regulate the individual genes within MNs on MEA plates, and monitor the neuronal excitability. To validate the effectiveness of this approach, we cultured purified MNs from the parental Hues3 *SOD1*$^{WT/WT}$ hES cell line with murine glia for 15 days, then used Lipofectamine RNAiMax to transfect the mature cultures with either a scramble siRNA, a scramble red fluorescentl (conjugated to Alexa Fluor 555 molecule) control siRNA, or an siRNA targeting KCNQ3 (Figure 3.14A). We chose KCNQ3 because it is a direct target of retigabine and was found to be expressed in our MNs. Three days after the treatment, we dissociated the cultures and used an NCAM antibody which was conjugated to a 647 fluorophore to purify the MNs away from the murine glia. FACS analysis confirmed the absence of NCAM labeling in a culture in which murine glia were plated without the addition of MNs, whereas cultures containing both glia and MNs had an NCAM positive fraction (Figure 3.14A). This result demonstrated the ability of NCAM labeling to identify MNs when previously purified from other neuronal cell types. FACS analysis revealed that over 95% of the NCAM-positive cells were co-labeled with the control red fluorescent siRNA, demonstrating the ability of the siRNAs to enter MNs in the mixed cultures. Furthermore, when qPCR was performed on the RNA isolated from the NCAM-positive fraction of the cultures which were treated with the KCNQ3 siRNA, the levels of KCNQ3 were found to be reduced to below 15% when compared to the scramble control (Figure 3.14B).

After confirming the efficiency of siRNA knockdown in our cultures, we designed an initial pilot study in which we would reduce the levels of KCNQ3 and SLC10A4, each gene

having been found to be expressed in our MNs and implicated in MN excitability. We generated MNs from the parental Hues3 *SOD1*$^{WT/WT}$ cell line using the adherent differentiation protocol, and used FACS to isolate MNs based on GFP fluorescence. After FACS purification we plated 50,000 MNs with 50,000 murine gla on a 12 well MEA plate. After 18 days of culture, we treated the cultures with a KCNQ3, SLC10A4, or scramble siRNA. To determine the impact of the transfection on spike rate, we withheld treatment of some wells for comparison with the scramble siRNA. We found that the application of the scramble siRNA resulted in a significant decrease in spike rate 10 days after transfection (Figure 3.14C). However, the considerable variability between wells suggested that many wells per treatment would be necessary for reliable results to be obtained. Unfortunately, this finding limited the utility of the 12-well MEA plates for future experiments, given the time and labor necessary to treat multiple wells across several plates, as would be required to investigate a large number of the transcripts identified as possible modulators of MN excitability. Thus, in order to optimize siRNA transfection during MEA experiments while maintaining many treated wells per condition, we established MN cultures in 96-well and 48-well plates, containing 8 and 16 electrodes per well respectively.

Figure 3.14

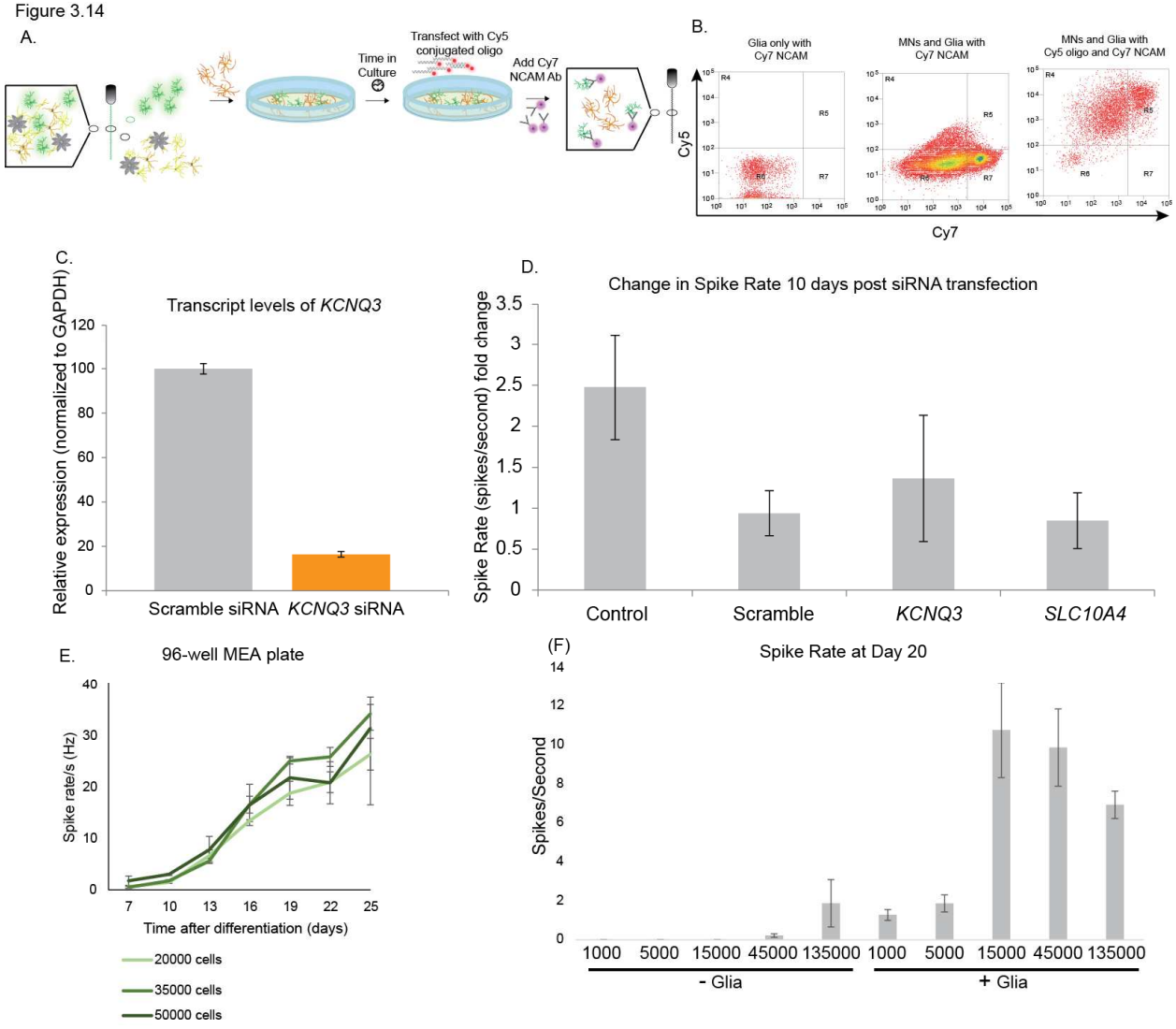

Figure 3.14. MEA recordings after reduction of target transcripts. (A) Diagram of experiment. (B) FACS analysis confirmed the transfection of NCAM-positive MNs with a flourescently conjugated oligonucleotide. (C) KCNQ3 levels could be dramatically reduced by siRNA transfection. (D) siRNA transfection causes a change in physiological activity when compared to unstransfected controls. (E) 96-well MEA plates provide reliable spike rate recordings. (F) Co-culture with murine glial cells has a dramatic impact on the physiology of MN cultures.

First, we performed a pilot experiment aimed at determining the appropriate number of cells required for robust spike rates in 48- and 96-well MEA plates. For both the 48- and 96-well MEA plates, we added 20K, 35K, or 50K MNs per well. At the same time as the MNs were plated, we also added 60K murine glia for the 48-well plate and 40K murine glia for the 96-well plate. To determine the spike rates in our cultures, we took MEA recordings every 3 days for 3 weeks. As a cutoff criterion for reliable measurements, we only used recordings from which 75%

of the electrodes had a spike activity above background. Using this criteria, we found a large number of sub-threshold wells in the 48-well plate when 20K, 35K, or 50K MNs were plated with 60K murine glia (Data not shown). We therefore eliminated the 48 well plate from further experiments. The 96 well MEA plate performed reliably with all conditions tested (Figure 3.14D).

After confirming the functionality of the 96-well MEA plates, we decided to use the larger format to address a question which had previously been difficult due to the lack of a larger multiwell MEA platform. We were concerned that variability in the percentage of MNs produced across differentiations would substantially influence the total number of MNs on the MEA plates, thereby creating noise and reducing the reproducibility of our findings. In order to determine the impact of varying plated MN numbers on MEA recordings, we first used the adherent differentiation protocol to generate MNs from the previously described 39b patient-specific iPSC cell line[9]. We then plated 1K, 5K, 15K, 45K and 135K unsorted cells per well in a 96-well MEA plate in the presence or absence of murine glia, and monitored the excitability of each well after 20 days. For the wells which contained murine glia we found very little difference in spike rate in wells which contained 1K or 5K cells per well (Figure 3.14E). However, the spike rate dramatically increased when the number of cells was increased to 15K per well. Thereafter, the spike rate remained fairly constant in wells containing between 15K and 135K cells per well. While these results were surprising, they demonstrated that the spike rate was not majorly impacted by the cell number within a range of 15K and 135K. Given this information, we decided to plate 50K cells per well in future experiments, to minimize the likelihood that cell number would drive differences in spike rates. Another interesting observation from these experiments was the dramatic increase in spike rates observed when the MNs were cultured with murine glial support. This observation demonstrates the importance of glial support, and sheds light on the transcriptional differences between the RNA sequencing experiments described in the current chapter and the previous study (described in chapter 2) using patient-specific iPSCs[9]. Ideally, MEA recordings from purified MNs in the absence of glia would

facilitate further exploration of the relationship between physiological activity and glial support. Unfortunately, we have not been able to obtain reliable recordings from purified MNs on the MEA plates without the addition of murine glia. We believe that this is due to the clumping of purified MNs in the absence of glia, which leads to heterogeneous coverage of the electrodes within a well, and therefore variable recordings when comparing multiple wells. Currently, we are testing multiple surface conditions to find a coating which will allow the MNs to mature in culture without clumping. We will then be able to determine the electrophysiological properties of the MNs in the absence of murine glial cell types.

Currently we are using the 96 well MEA plates to optimize siRNA transfection conditions. After optimal conditions are found we will proceed with knocking down the transcript levels of our chosen targets and assessing the impact on spike rate.

Discussion

Our previous findings established robust phenotypic differences in MNs expressing the *SOD1 A4V* variant when compared to controls[9]. Findings from gene editing experiments confirmed the necessity of the mutation for the modulation of pathways observed in culture and transcriptional analysis[9]. The aims of the current study were to determine the sufficiency of the *SOD1 A4V* mutation to drive phenotypic changes when introduced into a control hES genetic background, as well as to investigate the cell autonomous nature of the previously described phenotypes by performing experiments on cultures containing purified MNs. We were able to confirm a number of our previously demonstrated phenotypes using the purified MN cultures, and are currently working towards understanding discrepancies between these and previous results. Importantly, we have improved upon the methodologies used in previous experiments to develop more robust and high throughput assays.

The RNA sequencing experiment described here is, to our knowledge, the first attempt at discovering transcriptional changes which develop as stem cell derived MNs develop in culture. While our main goal was to uncover maturation-dependent genotypic differences, this

data set has also provided incredible insights into the impact of in vitro culture on MN development. It may be possible to use this data set to determine signaling pathways capable of driving MNs to more mature states in vitro, without the need for murine glial support. While our current data sets generated with purified MNs derived from the parental and gene target Hues3 cell lines do not show enrichment in mitochondrial dysfunction or ER stress, we do find some indications that these pathways are modulated in the cells. First, using EM we observed malformed mitochondria in MNs when analyzed at day 30. Second, increased XBP1 splicing was found in the *SOD1*$^{A4V/WT}$ MNs at later time points. These results indicate that the degenerative process is underway in the purified MNs, but that the timing may be delayed, possibly due to the absence of direct interaction with murine glial support. In sum, the current data set provided unique methodological advances and new information about the processes which influence the output of neurodegenerative disease modeling experiments, which promise to shape the future of stem cell disease modeling in important ways.

Our finding of increased excitability in purified MNs expressing the *SOD1 A4V* mutation confirms this pathway as a novel, therapeutically relevant phenotype. Using gene targeting and reporter cell lines, we established a well-controlled system amenable to high throughput screens. This will benefit future experiments by reducing the chance of indirect effects of small molecules on non-MNs in the culture and increases the chances of finding drugs which act directly on the MNs to alter physiological properties. After determining appropriate conditions for siRNA knockdown within MEA wells, we are now poised to uncover the impact of reducing several interesting candidates found in our RNA sequencing experiments. We are hopeful that these findings will lead to more therapeutically relevant targets, which can then be screened using more traditional, higher throughput experiments, or already have known small molecule modulators.

Methods

Cell Culture. Cell were cultured at 37C in the presence of 5% CO2. All pluripotent stem cell culture was performed with mTesr media with the use of matrigel coatings. hES and iPSCs were allowed to expand for 4-5 days and then passaged by incubation with 1mM EDTA for 5 minutes. The EDTA was then removed and 1 mL of mTesr was used to dislodge the cells from the plate. The cells were then passaged 1:10 into a freshly coated matrigel coated dish. All neurons were cultured with MN media (Neurobasal, Glutamax, NEAA, N2, and B27) with ROCK inhibitor, GDNF, BDNF, and CNTF. For glial conditioned media, MN media was conditioned in the presence of glia for 72 hours. The media was then filtered with a 0.45um filter and frozen at -80C. The glial conditioned media was then thawed and diluted 1:! with fresh MN media. BDNF, GDNF, and CNTF (10mg/mL each) were added fresh to the 1:1 mixture the day of feedings.

Gene Targeting. Zinc finger nucleases (ZFNs) were constructed using either the OPEN method as described previously (Maeder et al., 2008) or a modified version of OPEN that uses antibiotic resistance for the selection as previously described (Sander et al., 2011). Briefly, pools of ZF pre-selected zinc finger (ZF) domains were ligated together to create a combinatorial library of three-3 finger proteins. A bacterial two-hybrid-based selection system was used to interrogate the ZF library for proteins that could bind the appropriate target sequences of interest. ZF proteins that bound the target sequence were cloned into a mammalian expression vector and fused to heterodimeric FokI nuclease protein domains to construct ZFNs. Active ZFNs capable of inducing a double strand break at the desired locus were identified by screening pairs of nucleases for the capability to induce characteristic indel mutations at the SOD1 target site in sequencing the locus of ZFN-treated HEK 293s. 2.5 million iPS cells of the 39b cell line were accutased and nucleofected using Human Stem Cell Kit II and program A-023, with 1 mg of ZFN plasmid and 5mg of targeting plasmid. After nucleofection, the cells were plated on matrigel with mTesr and ROCK inhibitor. After 48 hours, puromycin selection was applied for 1 week, after which surviving colonies were passaged and gDNA was extracted.

PCR was used to confirm proper targeting of the cassette. To remove the puromycin cassette, 2.5 million cells were nucleofected with 1mg of a mammalian expression plasmid containing hygromycin and 5mg of a mammalian expression plasmid containing the FLPo recombinase. The cells were then plated on matrigel with mTeSR and ROCK inhibitor. Twenty four hours after nucleofection, hygromycin was added for 48 hours. Colonies were allowed to expand for 1 week, after which time they were picked and genomic DNA was extracted. Sequencing of the genomic DNA was used to confirm removal of the puromycin cassette. SOD1 expression was verified by qPCR after RNA extraction and cDNA synthesis. PshAI digestion along with sequencing of the qPCR product demonstrated loss of expression of the mutant allele. For CRISPR targeting, a gRNA targeting the SOD1 locus was cloned into plasmid pX335 (Addgene). This plasmid was substituted for the ZFN plasmid and all other steps were replicated as described above.

MN differentiation and purification

Genome Sequencing. Pluripotent stem cells were washed once with PBS, then scraped off the plates, resuspended in PBS, and spun down. The PBS was then aspirated off and the pellet was resuspended in cell lysis buffer. DNA was extracted using standard phenol chloroform methods. DNA sequencing libraries were generated using the Illumina Nextera DNA sample preparation kit. The libraries were sequenced on an Illumina HighSeq 2500.

MN Differentiation. EB motor neuron differentiation was carried at as previously described by our lab. Briefly, pluripotent stem cell colonies were dissociated with accutase and single cells were plated in suspension in low-adherence dishes, at a 400K/ml density with 10mM ROCK inhibitor (Sigma, Y-27632) in mTeSR media for 24hrs. Embryoid bodies (EBs) were formed and media was gradually diluted (50% on day 3 and 100% on day 4) to KOSR (DMEM/F12, 10% KOSR) between days 1-4 and to a neural induction medium (NIM: DMEM/F12 with L-glutamine, NEAA, Heparin (2mg/ml), N2 supplement (Invitrogen) for days 5-24. From days 1-6 cells were cultured in the presence of SB431542 (10mM, Sigma Aldrich) and

Dorsmorphin (1mM, Stemgent), and from days 5-24 with BDNF (10mg/ml, R&D), ascorbic acid (AA, 0.4mg/ml, Sigma), Retinoic Acid (RA, 0.1mM, Sigma) and Smoothened Agonist 1.3 (SAG 1.3, 1mM, Calbiochem). At day 24 EBs were dissociated to single cells with Papain/DNase (Worthington Bio) and plated onto poly-lysine laminin-coated chamber slides/plates/coverslips (BD Biosciences) for relevant experiments. For adherent differentiation of MNs the pluripotenent stem cells were dissociated to single cell with accutase and then 1,000,000 cells were plated in mTesr with ROCK inhibitor in one well of a 6-well plate. The cells were fed with mTesr until they reached confluency (typically 2 days) then the media was changed to 2D-Diff media (1:1 Neurobasal:DMEM,F12 with Glutamax, NEAA, N2, and B27 supplements) with the addition of SB431542 (10mM), LDN (100nM), RA (0.1mM), and SAG 1.3 (1mM) for days 0-5. On day 6 the media was changed to 2D-Diff media with RA (0.1mM), SAG 1.3 (1mM), DAPT (5mM), and SU4312 (5mM). On day 14 the cultures were dissociated with Accutase and plated in MN media (Neurobasal, Glutamax, NEAA, N2, and B27) with ROCK inhibitor, GDNF, BDNF, and CNTF. The next day the media was changed to MN media with GDNF, BDNF, and CNTF. The media was then refreshed every 3 days.

FACS Analysis. Day 24 EB differentations were dissociated with Pappain or day 14 Adherent differenations were dissociated with Accutase for forty five minutes. The cultures were then pipetted until the cells were single cells and spun down. The neurons were resuspended in MN media with ROCK inhibitor, GDNF, BDNF, and CNTF. The Neurons were then passed through a 0.45 um filter and then MN media was added (4 mLs for every well of dissociated cultures). The cells were then sorted for GFP fluorescence at the Bauer core. MNs were collected and plated in MN media with Rock inhibitor, GDNF, BDNF, and CNTF. The next day the MNs were fed with MN media containing GDNF, BDNF, and CNTF.

Immunocytochemistry. Cell cultures were fixed in 4% PFA for 15minutes at 4℃, permeabilized with 0.2% Triton-X in PBS for 45 minutes and blocked with 10% donkey serum in PBS-T (Triton 0.1%). Cells were then incubated in primary antibody overnight and secondary

antibodies for 1 hour in 2% donkey serum in PBS-T after several washes in between. DNA was visualized by a Hoechst stain. The following antibodies were used: Primary antibodies used in this study are Islet1 (1:200, DSHB, 40.2D6), TUJ1 (1:1000, Sigma, T2200), SOD1 (1:2000, Agrisera #AS09 540). Secondary antibodies used (488, 555, 594, and 647) were AlexaFluor (1:1000, Life Technologies) and DyLight (1:500, Jackson ImmunoResearch Laboratories).

Motor Neuron Survival Assay. After 24 days of differentiation, neuronal EBs were dissociated and FACS purified. 20K MNs were plated on poly-D-lysine/Laminin coated 8-well chamber slides (BD biosciences) containing a confluent monolayer of primary cortical mouse glia. Primary glial preparations from P0-P2 mouse pups were generated as described previously (Boulting et al., 2011; Di Giorgio et al., 2008). Fresh glial preparations (<1 month, <2 passages) were used. Co-cultures were maintained in Neurobasal media (NB, Invitrogen), supplemented with B27 and N2 supplement (Invitrogen), 10mg/mL each of BDNF, GDNF, CNTF (R&D) and fed every 3 days. Slides were fixed on days 3 and 30. Cultures were stained for TUJ! And Human Nuclei and cell numbers assessed. Whole-well images were quantified using imageJ. Neuronal numbers on day 3 were set as 100% and numbers day 30 were expressed as a percentage of day 3.

Mitochondrial EM Analysis. After 14 days of adherent differentiation, neuronal cultures were dissociated and FACS purified. 20K MNs were plated on poly-D-lysine/laminin-coated 35 mm glass coverslips (BD-Biosciences). MN cultures were fixed with 2.5% glutaraldehyde-2% paraformaldehyde in 0.1M sodium cacodylate buffer (pH 7.4) and maintained at 4℃ O/N. Cultures were then postfixed in 1% OsO4-1.5% KFeCN6 for 30 min, washed in water 3x and incubated in 1% aqueous uranyl acetate for 30mn followed by 2x washes in water and subsequent dehydration in grades of alcohol (5min each; 50%, 70%, 95%, 2x 100%). Cells were then embedded in plastic and ~60nm thick sections were cut, picked up onto copper grids, stained with lead citrate and analyzed in a JEOL 1200EX Transmission Electron Microscope. Pictures were taken by a technician blinded for sample IDs.

SOD1 Solubility. After 24 days of differentiation, neuronal EBs were dissociated and 500k cells were plated on poly-d-lysine and laminin coated 6-well plates. After 10 days of culture the wells were treated with 0.5 uM MG132 or DMSO control for 5 days. For analysis of SOD1 protein, detergent-soluble fractions were prepared using RIPA buffer and detergent-insoluble fractions were obtained using UREA buffer (8M UREA; 4% CHAPS; 40 mM Tris; 0.2% Bio-Lyte® 3/10 ampholyte). 5mg of detergent-soluble and equivalent volumes of detergent-insoluble protein samples were separated by SDS-PAGE (Bio Rad Laboratories), transferred to PDVF membranes and probed with anti-SOD1 antibody (Agrisera #AS09 540) and anti-a-Tubulin (Sigma Aldrich # T6199).

Sequential Time Point RNA Sequencing and Analysis. MNs were produced with the adherent differentiation protocol. On day 14 the differentiations were dissociated and MNs were purifed by FACS and 200,000 were plated in a 24-well plate coated with poly-d-lysine and laminin. The MNs were cultured in glia-conditioned MN media. Every other day from days 2 through 20, two wells of WT and two wells of SOD1 A4V MNs were washed once with PBS and then 250ul of Trizol was added to the wells. RNA was then extracted using standard purification methods. For next-generation RNA sequencing, RNA integrity numbers (RIN) above 7.5, determined by bioAnalyzer, were used for library preparation. In brief, RNA sequencing libraries were generated from ~250ng total RNA using the illumina TruSeq RNA kit v2, according to the manufacturer's directions. Libraries were sequenced at the Harvard Bauer Core Sequencing facility on a HiSeq 2000. All FASTQ files were analyzed using FastQC software (v 0.10.1) to confirm that Phred scores were acceptable at all read positions (median Phred score>25 and lower quartile>20). The FASTQ files were aligned to the GRCh37/hg19 reference genome using Tophat (v 2.0.7). Duplicated reads were removed using Picard Tools MarkDuplicates (v 1.44). Differential expression testing was performed independently using two separate analysis packages: Cufflinks (v 2.1.1) and DESeq. The Cufflinks output was visualized with the cummeRbund R package using a false discovery rate of 0.05. Gene Set Enrichment Analysis

(GSEA, Broad Institute) was performed by first creating a pre-ranked gene list of all genes included in differential expression testing ordered by log2 fold change. Analysis was performed using the GSEA preranked tool with the REACTOME and KEGG Pathway MSigDB collections.

XBP1 Splicing Assay. 300ng of RNA was used to generate cDNA. PCR reactions were set up using 2ul of cDNA and premixed *Ampligold Taq* Polymerase (Applied. Biosystems), at 66^{C} annealing. PCR products were analyzed on a Tapestation 2200 and analyzed with the Tapestation software.

RT-PCR. On day 14, adherent MN differentiation cultures were dissociated and purified by FACS. 200,000 sorted MNs were plated in poly-d-lysine and laminin coated plates. On day 20, RNA was isolated from the MNs using Trizol. Iscript cDNA synthesis kit was used to generate cDNA from the purified RNA. qPCR was performed using SYBER green.

MEA Recording. Differentiated MN cultures were dissociated and 50,000 sorted or unsorted MNs were combined with 50,000 murine glia. The MN and glia cells were spotted onto the 12 well MEA plates and allowed to attach for 30 minutes. After 30 minutes, media was added to all wells of the MEA plate. For MEA recording, the AXION MEA system was heated to 37 degrees. Subsequently, the MEA plate which was to be recorded was taken from the tissue culture incubator and placed in the AXION device for 1 minute. Recordings were then taken for 2 minutes. For 48-well and 96-well MEA experiments the cells were not spotted onto the electrodes. Instead, 200ul or 100ul of cells and glia mixture was added to the 48-well or 96-well plates respectively which would coat the entire surface area of the plates.

References

1 Sandoe, J. & Eggan, K. Human Neurodegenerative Disease Models: The Opportunities and Challenges Presented by Pluripotent Stem Cells. *Nature neuroscience* **In Review** (2013).

2 Andersen, P. M. & Al-Chalabi, A. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nature Reviews Neurology* **7**, 603-615 (2011).

3 Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59-62 (1993).

4 Gurney, M. E. *et al.* Motor neuron degeneration in mice that express a human Cu, Zn superoxide dismutase mutation. *SCIENCE-NEW YORK THEN WASHINGTON-*, 1772-1772 (1994).

5 Di Giorgio, F., Carrasco, M., Siao, M., Maniatis, T. & Eggan, K. Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model. *Nature neuroscience* **10**, 608-614, doi:10.1038/nn1885 (2007).

6 Oh, Y. K., Shin, K. S., Yuan, J. & Kang, S. J. Superoxide dismutase 1 mutants related to amyotrophic lateral sclerosis induce endoplasmic stress in neuro2a cells. *J Neurochem* **104**, 993-1005, doi:10.1111/j.1471-4159.2007.05053.x (2008).

7 Deitch, J. S. *et al.* Phenotype of Transgenic Mice Carrying a Very Low Copy Number of the Mutant Human G93A Superoxide Dismutase-1 Gene Associated with Amyotrophic Lateral Sclerosis. *PLoS ONE* **9**, e99879 (2014).

8 Gladman, M., Cudkowicz, M. & Zinman, L. Enhancing clinical trials in neurodegenerative disorders: lessons from amyotrophic lateral sclerosis. *Curr Opin Neurol* **25**, 735-742, doi:10.1097/WCO.0b013e32835a309d (2012).

9 Kiskinis, E. *et al.* Pathways Disrupted in Human ALS Motor Neurons Identified through Genetic Correction of Mutant SOD1. *Cell Stem Cell*, doi:10.1016/j.stem.2014.03.004 (2014).

10 Cowan, C. A. *et al.* Derivation of embryonic stem-cell lines from human blastocysts. *New England Journal of Medicine* **350**, 1353-1356 (2004).

11 Marchetto, M. C. N. *et al.* Non-Cell-Autonomous Effect of Human SOD1< sup> G37R</sup> Astrocytes on Motor Neurons Derived from Human Embryonic Stem Cells. *Cell Stem Cell* **3**, 649-657 (2008).

12 Lobsiger, C. S. & Cleveland, D. W. Glial cells as intrinsic components of non-cell-autonomous neurodegenerative disease. *Nat Neurosci* **10**, 1355-1360 (2007).

13 Yuan, S. *et al.* Cell-surface marker signatures for the isolation of neural stem cells, glia and neurons derived from human pluripotent stem cells. *PLoS One* **6**, doi:10.1371/journal.pone.0017540 (2011).

14 Di Giorgio, F. P., Boulting, G. L., Bobrowicz, S. & Eggan, K. C. Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-

causing mutation. *Cell Stem Cell* **3**, 637-648, doi:S1934-5909(08)00522-5 [pii] 10.1016/j.stem.2008.09.017 [doi] (2008).

15 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).

16 Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology* **31**, 822-826 (2013).

17 Shen, B. *et al.* Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nature methods* **11**, 399-402 (2014).

18 Boulting, G. L. *et al.* A functionally characterized test set of human induced pluripotent stem cells. *Nature biotechnology* **29**, 279-286 (2011).

19 Chambers, S. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature biotechnology* **27**, 275-280, doi:10.1038/nbt.1529 (2009).

20 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

21 Chambers, S. *et al.* Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nature biotechnology* **30**, 715-720, doi:10.1038/nbt.2249 (2012).

22 Wainger, B. *et al.* Intrinsic Membrane Hyperexcitability of ALS Patient-Derived Motor Neurons. *Cell Reports* (2014).

23 Ron, D. & Walter, P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol* **8**, 519-529, doi:10.1038/nrm2199 (2007).

Chapter 4: Implications and Future Directions

The studies included in this thesis have demonstrated the capability of modeling ALS with neurons derived from pluripotent stem cells. In addition to the studies described herein, several reports of disease modeling experiments for other late onset neurodegenerative disorders have been published[1,2]. These successful experiments have proven the utility of pluripotent stem cell technology for disease modeling, and the potential of this technology for the development of promising therapeutic compounds. While the sophistication of disease modeling experiments has improved vastly in recent years, there remain some important and concrete next steps toward a more complete understanding of disease mechanisms. This chapter will describe methodological "best practices" which I believe should be employed in future disease modeling experiments as well as future experiments which could address of the questions which arise from the results presented in the previous sections.

Minimizing the Presence and Impact of Off-Target Mutations

The inherent genomic variation of each pluripotent stem cell line, and the potential confounding element it presents, is a major source of concern in any disease modeling experiment[3]. Initial disease modeling experiments were performed without proper controls for this variability[2], likely due to technical limitations in the generation of ideal controls. However, advances in gene targeting technology have now routinized the generation of genetically corrected cell lines, thus eliminating much of the variability derived from individual differences in genetic backgrounds, and allowing for the illustration of causative relationships between mutations and phenotypes[4,5]. While these advances have been transformative, as with any new methodological advancement, care must be taken to ensure that the conclusions drawn from experiments using genetically manipulated cell lines are as accurate and free from bias as possible. Specific methodological concerns should be considered when planning such experiments, which will be described in detail below.

The primary consideration which must be deliberated when initiating a gene targeting experiment is which technology should be used to drive homologous recombination (HR) at the

chosen genomic loci. The uniquely troublesome nature of gene targeting in human pluripotent stem cells has resulted in the advent of many approaches aimed at introducing and removing nucleotides in ES or iPS cells. This thesis has described two such technique, one of which was the use of zinc finger nuclease (ZFN) technology to introduce a double strand break to drive HR[6]. However, developing functional ZFNs is technically challenging and laborious[7]. Recently, the CRISPR/Cas9 system has been shown to introduce site-specific double strand breaks with high accuracy and precision[8]. Because of the relative ease of use of the CRISPR/Cas9 system, it has become rapidly adopted. Thus, future gene targeting experiments will likely benefit from the less technically challenging CRISPR/Cas9 system.

While the fidelity of site specific nucleases has been shown to be remarkably high[9], using nuclease techniques will always carry the worry of unintentional mutations driven by non-specific activity[10]. The current thesis has described specific steps which may be used to overcome this problem. Firstly, off-target mutations have been shown to occur at sites which harbor some sequence similarity to the intended on-target site[10]. Therefore, mutations induced by nucleases which target different genomic sequences should have different off-target possibilities[10]. This non-random pattern of inducing off-target mutations indirectly provides a means for addressing the potential variability such mutations present. Specifically, by employing at least two independent nucleases to target sequences sharing little homology, scientists can generate independent gene-targeted clones which should not share any off-target sites. By confirming that the phenotype of interest is present in each clone, one can be sure that an off-target mutation is not driving the main finding. Second, mutations which have been corrected with gene targeting can subsequently be reverted to the parental genotype using another round of nuclease-driven HR. Any off-target mutations which were induced in the initial round of gene targeting should remain in the reverted cell line, and in this way, the reverted cell line can be used to confirm that the other off-target sites are not driving any phenotypes found. Third, whole-genome sequencing can be utilized to reveal any major aberrations within cell lines, in

order to exclude such lines from gene targeting experiments[9]. Lastly, the confirmation of results using non-isogenic patient-specific iPSCs, but which harbor the disease mutations, will provide more evidence that the results in question are not driven by-off site mutations[5]. Carefully employing the techniques described above should greatly reduce the likelihood of conclusions about disease mechanisms being driven by unintentional off-target mutations.

Addressing Culture-induced Mutations

In addition to off-site mutations introduced by gene targeting, additional sources of variation should be considered and addressed when conducting any experiment which requires the clonal isolation and subsequent propagation of a pluripotent stem cell line. For example, recent publications using whole-genome sequencing have demonstrated unavoidable accumulation of mutations as a result of in vitro culture[5,9]. Consistent with these findings, we too found the accumulation of SNVs in our gene targeted cell lines when compared with the parental cell line's genome[5]. These SNVs were likely spontaneously induced during in vitro culture, as they were not found near sequences which contained any homology to the intended nuclease site[5]. Unfortunately, current gene targeting technology necessitates clonal isolation steps which consistently result in a cell line not wholly isogenic to the parental line. As pluripotent stem cells are mitotically active, mutations will naturally accrue through various mechanisms[9]. Thus, a population of pluripotent stem cells which are descendants of a single cell derived from an embryo (for hESCs) or fibroblast (for iPSCs) will be a heterogenous mixture of cells with random SNVs derived from cell culture. Therefore, any cell isolated from the heterogeneous population will go through a genetic bottleneck, and necessarily the resulting cell line will have a slightly different genetic background when compared to the bulk parental population[9]. There are several steps, described in detail below, which can be taken to reduce the effects of culture-incurred mutations propagated by clonal isolation steps. First, the use of parental cell lines which are low passage or recently clonally isolated will reduce the genetic heterogeneity within the cells to be targeted[11]. Second, recent advances in gene targeting

technology have allowed for efficient single-step targeting using oligos[8]. This technique eliminates a second clonal isolation step, and will therefore reduce the number of culture-induced mutations between the parental and final gene targeted cell line. When oligo-based targeting is not possible and the use of an antibiotic selection cassette is necessary (e.g., when adding larger cassettes such as disease-associated expansions or fluorescent proteins for reporter cell line generation), it will be important to limit the number of passages before the selection cassette is excised from the intermediate cell line. This can be achieved by optimizing the selection cassette removal step so that very few cells are necessary for the removal step, and therefore fewer cell divisions are required. Finally, if multiple genomic sites are to be modified in a single cell line, advances in CRISPR/Cas9 technology now make possible multi-site targeting, allowing for multiple targeted modifications in a single gene targeting experiment[12]. In sum, these techniques serve to minimize culture-induced mutations by reducing the number of clonal isolation steps as much as possible without compromising the integrity of the experiment. Keeping this strategy in mind as a priority in future gene targeting experiments should serve to minimize deviations from isogenic cell lines, and thus maximize precision of results.

Identifying and Addressing Epigenetic Changes

While the majority of concerns about gene targeting technology have focused on off-site and culture-induced mutations[10], an often-overlooked but potentially as concerning product of gene targeting may be epigenetic variability. Studies in yeast have demonstrated the capacity for mammalian cells to retain small populations of epigenetically distinct progeny, which arise spontaneously and can be stably propagated after clonal isolation[13]. Ostensibly, a similar mechanism could result in the isolation of a gene-targeted cell line in which certain loci have been epigenetically silenced, or are more accessible than in the parental cell line from which it was derived. This could present a considerable problem for disease modeling experiments, if the altered loci had any bearing on disease phenotypes. Using existing technology, it is

conceivable to confirm the presence of or absence of epigenetic variability between parental and gene-targeted cell lines. For example, the same cell lines used for whole-genome sequencing in the current thesis could be used to acquire whole methylome data and examine the possibility of any variation in methylation sites. If any variation was identified, it may be possible to remove the epigenetic variability by exposing the parental and gene targeted progeny to epigenetic modifying small molecules. For example, treating aberrantly methylated cell lines with the DNA methyltransferase inhibitor 5-aza-cytidine (AZA) could remove the altered methylation and "reset" the cells lines to a common basal state of methylation[14]. Furthermore, the same "best practices" described above to minimize culture-induced genetic mutations should also reduce epigenetic variability in gene-targeted cell lines.

Modifying Mutations in Genetic Background

While ensuring that off-site mutations do not have an impact on results is a critical step for disease modeling experiments, there are other steps which should be taken in order to ensure robust and disease-relevant results. Most diseases, ALS included, have variable times of onset and rates of progression, even amongst patients harboring the same predisposing mutation[15]. While this variability likely has a large environmental component, studies have demonstrated that other modifying mutations are major drivers of variable onset and progression[15]. To circumnavigate this potential problem we used gene targeting to ensure an identical genetic background between the cell lines harboring the control and familial mutations[5]. In doing so, it is possible our results are biased if strong modifying mutations were present in the individual from whom the cell lines were drawn. We induced a familial ALS mutation into a control genetic background in the hopes we could identify robust phenotypes which did not require other modifying mutations for their identification. While we made an initial attempt at resolving this issue, it will likely require the use of many cell lines which have gene targeted counterparts from which to identify the impact of modifying mutations on a phenotype of interest. This approach would ensure any results driven by strong modifying mutations in a

single individual will be tempered by other parental and gene targeted pairs which lack the modifying mutations.

The use of large cohorts of parental and gene targeted pairs may also be helpful for grouping patients into meaningful phenotypically distinct populations as well as discovering single nucleotide variants (SNVs) which strongly influence disease progression in human cells. As described in the introduction, ALS has a considerable amount of phenotypic heterogeneity[15]. Utilization of whole genome sequencing technologies along with phenotyping large cohorts of patient-specific iPSCs and their gene corrected progeny could allow for the sub categorization of patients based on the performance of their iPSC derived MNs in disease modeling experiments as well as identify modifying mutations influencing cellular performance. For example, establishing a cohort of a hundred iPSC lines from patients harboring the same SOD1 A4V mutation, correcting the mutation in each cell line, and examining associations between the resulting phenotypes and the different modifying mutations present in each line may allow for improved understanding of the precise impact of specific mutations on disease progression. In addition, by simultaneously following the clinical outcome of each patient from whom the iPSCs were derived, it would be possible to make associations between in-vitro phenotypes and clinical outcomes. However, automation of the entire disease modeling process, from pluripotent stem cell derivation and culture to neural differentiation and phenotyping, will likely be necessary before a large enough cohort could be analyzed to yield convincing information about modifying loci. One approach which may minimize the number of cell lines needed to obtain such information may be to select iPSC lines derived from individuals with extreme clinical outcomes, thus increasing the likelihood of obtaining lines containing substantially influential modifying SNVs. If, in these cases, clinical progression of the patient was mirrored by performance in in-vitro experiments, this would represent strong evidence for the presence of modifying SNVs.

How Generalizable are the SOD1 Findings to other Familial Forms of ALS?

Gene targeting provides a unique tool for isolating and understanding the impact of single disease variants, by generating nearly isogenic cell lines except for the single variant being studied[5]. While this kind of knowledge will provide important pieces of the puzzle in our quest to more fully understand the etiology and pathophysiology of neurodegenerative diseases, a broader approach will be needed in order to understand how these puzzle pieces come together in a larger picture. This broader approach will likely require examination of cells derived from a large number of patients harboring different disease-causing mutations. For example, there are over 20 known familial mutations for ALS, which comprise almost the entirety of familial cases[15]. Taking a gene targeting approach to each of these mutations will allow for the generation of patient-specific iPSCs harboring each mutation, and their subsequent corrected control cell line. Transcriptional profiling and phenotyping of such genetically controlled pairs will illuminate any pathways which show signs of degeneration and are shared amongst the different familial mutations. These pathways may be more therapeutically advantageous, as they would likely be efficacious in a larger number of patients. Furthermore, as most of the ALS cases are idiopathic in nature, understanding the common pathways in a majority of the familial cases will likely be the most promising path toward developing therapies for sporadic cases[15]. The current thesis has described initial efforts in this vein, by using qPCR to identify common transcriptional changes between SOD1 and C9orf72 mutant MNs and MEA recordings to demonstrate increased excitability in SOD1, FUS and C9orf72 mutant MNs. Future studies should aim to continue identifying such common disease pathways, and work toward the development of therapeutics to normalize these disease phenotypes.

One specific approach to this goal may be to utilize gene targeting to introduce each familial mutation into a control ES cell background. Using this methodology, a single parental cell line would provide a control for the larger mutant cohort, eliminating the confounding influence of background variability between pairs of isogenic cell lines. Common molecular perturbations found in the cell lines harboring the familial mutations would represent an exciting

avenue toward understanding common mechanisms of ALS progression across disease-causing mutations. Additionally, to gain further insight into ALS pathology one could introduce familial mutations found in Parkinson's or Alzheimer's cases into the control cell line. Common molecular and cellular pathways found to be perturbed from the parental control cell line only amongst the cells harboring familial ALS mutations, and not Parkinson's or Alzheimer's mutations, would be most interesting to pursue.

How Best to Generate Meaningful RNA Sequencing Data

The overarching goal of the studies described in this thesis was to find genes or pathways necessary for ALS disease progression. We successfully identified multiple pathways which were altered by a familial SOD1 mutation, leading to the identification of retigabine as a novel possible therapeutic drug[16]. Still, we were unable to identify single genes which were modifiers of the disease. While we did find a large number of genes which were differentially expressed between disease and control MNs, we found it complicated to filter out transcriptional changes which are important for disease pathology. This complication resulted in a change of direction within the current project, our focus shifted increasingly onto the pathways which were identified, rather than specific genes.

If future studies are to build on our findings, it will be important to carefully select controls which can be transcriptionally profiled, and employ these to filter out transcriptional changes not likely necessary for disease manifestation. One possible such method is to utilize cell types known to be disease-resistant in vivo. As described in the introduction, sensory neurons are resistant to ALS, and importantly protocols exist to generate these neurons from pluripotent stem cells[17]. Future studies using isogenic lower MN and sensory neurons harboring the WT or A4V SOD1 allele for transcriptional profiling would allow for removal of many genes not relevant to disease pathology, and highlight clinically relevant disease mechanisms. Another interesting means for generating meaningful transcriptional data may be to treat MNs with small molecules known to exacerbate or rescue phenotypes, and then perform RNA sequencing. By looking for

gene expression differences which are markedly increased when the MNs are treated with the exacerbating drugs, and significantly decreased when small molecules known to rescue in vitro phenotypes are applied, specific genes relevant to disease progression may be highlighted. For example, we found that retigabine could rescue the hyper excitability and cell survival phenotypes found in MNs derived from the 39b iPSC line. Treating these MNs with retigabine, and then comparing their transcriptional signature with MNs derived from the 39b iPSC cell line which have not been treated with retigabine as well as the genetically corrected control MN treated with retigabine, could identify important transcripts which are modulated and lead to rescue of the survival deficit.

The expression profiling experiments described in this thesis provided a panoramic view of the in vitro disease course. A major disadvantage of our approach was the population-level transcriptome we obtained. While we undertook strategies aimed at obtaining a more homogeneous cell population, the use of a single gene is often not rigorous enough for creating a completely identical cell cohort[17]. Therefore, some of the changes observed in our transcriptional data may reflect an alteration in the various MN subtypes found within the population of cells being studies, rather than the up- or down-regulation of a gene across all cells in the dish. This issue is particularly important for the study of neurodegenerative disorders, since cell subtype specificity is often a key pathological finding[17]. Development of more sophisticated reporter cell lines which can parse MN subpopulations more finely is one promising avenue for addressing this issue. Additionally, recent advances in single-cell RNA sequencing technology may allow for identification of MN subtypes and determination of possible disease signatures simultaneously[18]. As the molecular underpinnings of MN subtype identity become better understood, this transcriptional code will allow for the prescribing of single MNs into the correct subtype class after sequencing[18]. After sub-categorizing the bulk MN population, it will be possible to identify signatures indicative of increased susceptibility. Furthermore, by using single cell RNA sequencing at multiple time points, it will be possible to

determine the dynamic nature of a purified MN population, and possibly identify susceptible subpopulations based on the loss of cell-type signatures[18].

Generating Mature Neurons for Disease Modeling Experiments

Possibly the most important complication of modeling late onset neurodegenerative disorders is the unknown factor of in vitro maturation and aging. Traditionally, chronological age has been used to describe the maturation state of neuronal cultures. The studies described herein used a range of time points from days 2 to 30. Other studies have used time points as late as day 75[19]. Highlighting the centrality of this issue, the results of the experiments described in Chapters two and three of the current thesis implicate the maturation state of the MNs as potential drivers of phenotypic inconsistencies between experiments. Specifically, when comparing the RNA sequencing experiments from Chapters two and three, there was very little overlap between significantly altered genes. While we present multiple rationales to explain these inconsistencies, I believe the presence of direct contact between the purified MNs and murine glial cell types in one data set (see Chapter 2) but not the other (see Chapter 3) is the most likely cause of the discrepant results. Two main findings have led me to this hypothesis. First, we have demonstrated through MEA recordings that a dramatic increase in electrophysiological activity occurs when neuronal cultures are plated on glial monolayers. Second, we have established a correspondence between neuronal firing rates and ER stress. These findings indicate that the MNs which were cultured in the absence of a glial monolayer likely had lower levels of physiological activity. In turn, this may have led to a reduced level of ER stress, which is known to cause widespread transcriptional and translational alterations[20]. It would therefore be no surprise to find differences in the significant genes identified between experiments, if the cells being analyzed had considerable differences in their ER stress levels due to culture conditions. Currently, we are investigating this hypothesis by performing the transcriptional profiling experiment with the parental Hues3 $SOD1^{WT/WT}$ and Hues3 $SOD1^{A4V/WT}$ ZFN_1 in with the same methodology described in Chapter 2. An alternate approach would be

to integrate an expression construct carrying an epitope-tagged ribosomal subunit into the genome of the parental Hues3 SOD1$^{WT/WT}$ and Hues3 SOD1$^{A4V/WT}$ ZFN_1[21]. The MNs could then be purified by FACS directly after differentiation and plated down on a glial monolayer. After 15 days of culture, the cells could be lysed and transcripts which are actively being translated could be pulled down by immunoprecipitating the epitope-tagged ribosomal subunit[21]. This would allow for purification of the human MN RNA for RNA sequencing libraries, but not necessitate culturing MNs in isolation.

Data from RNA sequencing experiments performed at multiple sequential time points may also serve as a resource for future experiments aimed at artificially increasing the maturation state of neurons. With this data at our disposal, it may be possible to introduce genetic factors necessary for neuronal maturation, to reduce the time necessary to produce the mature cells most vulnerable to disease processes. We found the Inhibitor of DNA binding (ID) class of transcriptional repressors to be dramatically reduced during the first 10 days of MN culture (Figure 4.1). It is possible that using siRNA to reduce the levels of the ID transcripts could hasten MN differentiation. Furthermore, the transcript levels of the POU3 class of transcription factors increased during the first two weeks of MN culture, implicating these transcription factors in the maturation of the MNs in vitro (Figure 4.2). Future studies which overexpress or reduce the levels of the factors found to be differentially expressed as the MNs mature may uncover factors which can decrease the time necessary for the production of mature MNs in vitro.

Figure 4.1

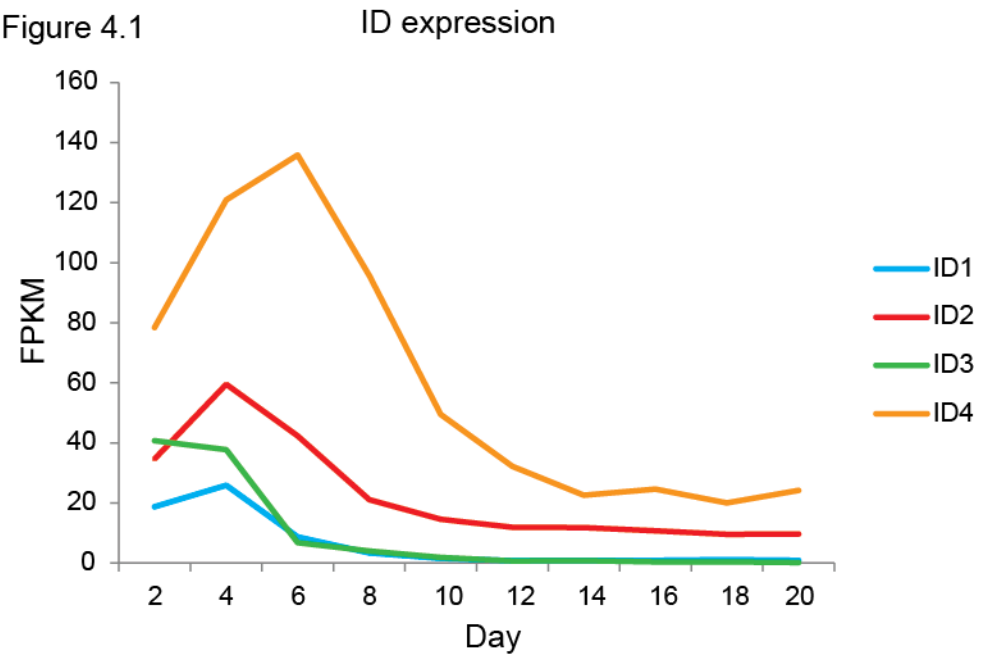

Figure 4.1 The Inhibitor of DNA Binding family of transcription factors were found to be downregulated in MNs over time.

Figure 4.2

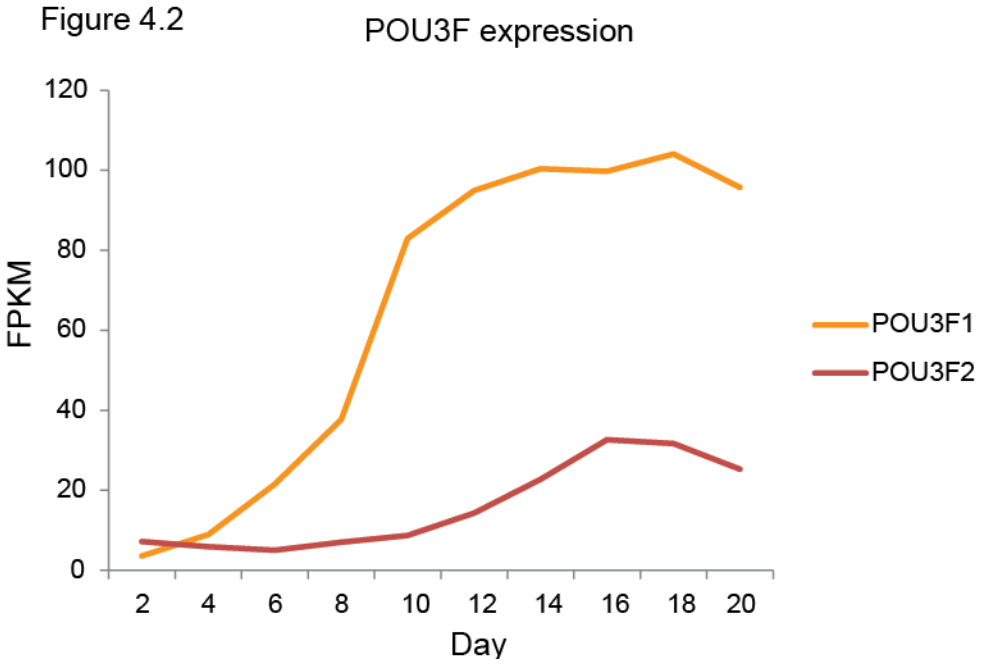

Figure 4.2. POU3F1 and POU3F2 were found to be upregulated in MNs over time.

How can Culture Conditions Develop a Broader Insight to ALS?

The experiments described in this thesis have focused on the cell-autonomous impact of

familial ALS mutations. Understanding of the molecular changes driven by the mutant SOD1 protein in MNs alone is an important step in uncovering the underpinnings of degeneration in ALS. However it has been previously shown that many other cell types are involved in the disease course[22,23]. Use of stem cell-derived cell types for modeling complex cell to cell interactions could enable dissection of the complex interplay between the various cell types involved in ALS.

Many reports have demonstrated that astrocytes and microglia expressing familial mutations are toxic to MNs[23,24]. Most of these studies have involved primary rodent tissue and have focused on SOD1 mutations, as rodent models overexpressing various SOD1 alleles are readily available[24]. These studies of toxic non-cell autonomous astrocytes and microglia have implicated many possible therapeutic pathways, such as inflammation and the immune response[25]. However, many of these models suffer from the use of higher-than-normal expression of the mutant protein, as well as being from rodent models, which have not provided good predictive therapeutic value when compared to human trials in the past[26]. The use of pluripotent stem cell-derived astrocytes and microglia to explore the non-cell autonomous impact of familial ALS mutations on MNs would be an important next step for disease modeling of ALS. Currently, this avenue of research is limited by several factors. First, there are far fewer protocols for the differentiation of astrocytes, and fewer still for microglia[27,28]. Many of the current astrocyte differentiation protocols require a prohibitively long period of time and have not been widely used[27,28], leaving their replicability in question. Second, while the molecular road map of MN development has been intensely studied for many years[29,30], the transcription factors and other cellular markers leading to the establishment of astrocytes are less well-studied. This creates complications in generating astrocytes from stem cells, as there is not a widely-accepted marker which can be used to delineate astrocytes. Third, the maturation of astrocytes in vitro may be more complicated than that of MNs, and may in fact require other cell types, leading to an even more complex in vitro system[31]. These problems will need to be overcome

before more complex disease modeling experiments, combining MNs, astrocytes, and microglia, can be conducted.

The interaction between MNs and other glial cell types is important and will provide valuable insight into disease mechanisms. However, perhaps the most important interaction to model in vitro is the interplay between MNs and muscle. The breakdown of the motor unit has been shown to be one of the first pathologies evident in rodent ALS models[32,33]. Additionally, experiments which increase the re-innervation of muscle by surviving MNs have been found to increase the lifespan of the ALS mouse model[33]. These findings indicate that the MN and muscle interaction could be a very therapeutically valuable system to model in vitro. Our lab and others have used primary mouse muscle to demonstrate the ability of our stem cell-derived MNs to innervate and form neuromuscular junctions (NMJs) in vitro[34,35]. However, these experiments have proven extremely laborious, and their results have been variable. A major impediment to routine MN/muscle co-culture systems is the need to derive muscle from primary tissue. Recently, stem cell differentiation protocols for generating muscle have been described, possibly providing an alternative to the use of primary cells[36]. Ultimately the reconstruction of the NMJ may be the most important in vitro model which can be established for diseases in which spinal MNs are involved.

While this report has focused on the impact of familial ALS mutations on spinal MNs, the disease is not exclusive to these neurons. Upper MNs also degenerate during ALS disease progression, and should be taken into consideration when establishing ALS disease models[15]. The major obstacle impeding widespread interrogation of upper motor neurons in stem cell disease modeling experiments has been the lack of available differentiation protocols. However, protocols for the generation of cortical MNs are readily being produced, and will hopefully provide this much needed cell type.

In light of the bright future unfolding with respect to stem cell differentiation protocols, it may one day be possible to recapitulate the entire in vivo cellular environment from stem cells.

This could allow for intricate studies involving the mixing and matching of patient and gene-corrected cell types, which could provide a window into what cell types are causing the different pathologies commonly seen in patients. Furthermore, the establishment of the complex circuit between cortical MN, lower MN, muscle, and possibly sensory neurons will allow for very informative physiological experiments to be conducted. It is likely that this complex culture may require advances not only in current differentiation protocols, but also in tissue culture techniques. In addition, it is important to remember that plating all the necessary ingredients down into a dish is not equivalent to establishing the connections and interactions which are important in vivo. Most likely, the use of microfluidic devices will be needed to facilitate the establishment of the correct circuits and also to provide a more convenient way to monitor these interactions. For example, astrocytes, microglia, and upper MNs could be plated in one chamber, which would allow the axons of the upper MNs to pass into a chamber containing astrocytes, microglia and lower MNs, which in turn would provide an avenue for the lower MN axons to access a chamber containing muscle. Various tracing dyes or viruses could be used to follow the connections made in the device, and more importantly, study the degeneration of these connections as disease processes progress.

Translation of Findings to Clinical Trials

Previous studies employing mouse models of ALS have thus far failed to produce efficacious therapeutic compounds[37]. This thesis has described the use of pluripotent stem cells to establish human models of the disorder, and importantly, discovered the FDA approved drug retigabine as a potential modifier of disease pathology. The utility of human stem cell-based models will ultimately be decided in the clinic, as therapeutic compounds discovered using these techniques are tested in human trials. In addition to the discovery of therapeutic compounds, use of patient-derived stem cells will also shed light on the most effective methods for undertaking human trials. As described earlier, segregating patients based on in vitro phenotypes may allow for the re-classification of patients into subgroups, which may have

different manifestations of the disorder[15]. This data could, in turn, provide a path for the establishment of more efficient and predictive clinical trials. For example, if small molecules are found which only rescue phenotypes only in patient-derived iPSCs which harbor SOD1 fALS mutations and not C9orf72 or TDP43, then clinicians could design a trial including only patients with SOD1 mutations, rather than including all ALS patients in the trial and thus diluting sub-group therapeutic benefits.

Conclusions

In this chapter, I have described my suggestions for next-step studies in the area of stem cell modeling of neurodegenerative diseases. While not comprehensive, each step may be a promising avenue toward improved understanding of disease etiology and the development of therapeutics, particularly given the rapidly advancing technology and protocols available today. More importantly, they can provide evidence for or against the limits of in vitro disease modeling with stem cell derived neurons. The use of a common genetic background in which to introduce known familial mutations will allow for comprehensive insight into the genetic influence of neurodegenerative disease, which could not be obtained using other currently available methods. Furthermore, comparing gene targeted cell lines from large cohorts of individuals with the same mutations may help to explain the heterogeneity in clinical presentations observed amongst patients. Additionally, carefully conducting transcriptional profiling experiments with well thought out controls, thereby highlighting clinically relevant alterations, will point toward genes for which overexpression or reduction may modify disease course. In sum, these next-step experiments will contribute to improved understanding of ALS and the use of human pluripotent disease models.

References

1 Han, S. S. W., Williams, L. A. & Eggan, K. C. Constructing and deconstructing stem cell models of neurological disease. *Neuron* **70**, 626-644 (2011).

2 Sandoe, J. & Eggan, K. Human Neurodegenerative Disease Models: The Opportunities and Challenges Presented by Pluripotent Stem Cells. *Nature neuroscience* **In Review** (2013).

3 Bock, C. *et al.* Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* **144**, 439-452 (2011).

4 Ding, Q. *et al.* Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* **12**, 393 (2013).

5 Kiskinis, E. *et al.* Pathways Disrupted in Human ALS Motor Neurons Identified through Genetic Correction of Mutant SOD1. *Cell Stem Cell*, doi:10.1016/j.stem.2014.03.004 (2014).

6 Maeder, M. *et al.* Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Molecular cell* **31**, 294-301, doi:10.1016/j.molcel.2008.06.016 (2008).

7 Maeder, M., Thibodeau-Beganny, S., Sander, J., Voytas, D. & Joung, J. Oligomerized pool engineering (OPEN): an 'open-source' protocol for making customized zinc-finger arrays. *Nature protocols* **4**, 1471-1501, doi:10.1038/nprot.2009.98 (2009).

8 Shen, B. *et al.* Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nature methods* **11**, 399-402 (2014).

9 Veres, A. *et al.* Low Incidence of Off-Target Mutations in Individual CRISPR-Cas9 and TALEN Targeted Human Stem Cell Clones Detected by Whole-Genome Sequencing. *Cell Stem Cell* **15**, 27-30 (2014).

10 Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology* **31**, 822-826 (2013).

11 Suzuki, K. *et al.* Targeted Gene Correction Minimally Impacts Whole-Genome Mutational Load in Human-Disease-Specific Induced Pluripotent Stem Cell Clones. *Cell Stem Cell* **15**, 31-36 (2014).

12 Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-918 (2013).

13 Grossniklaus, U., Kelly, W. G., Ferguson-Smith, A. C., Pembrey, M. & Lindquist, S. Transgenerational epigenetic inheritance: how important is it? *Nature Reviews Genetics* **14**, 228-235 (2013).

14 Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49-55 (2008).

15 Andersen, P. M. & Al-Chalabi, A. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nature Reviews Neurology* **7**, 603-615 (2011).

16 Wainger, B. *et al.* Intrinsic Membrane Hyperexcitability of ALS Patient-Derived Motor Neurons. *Cell Reports* (2014).

17 Kanning, K. C., Kaplan, A. & Henderson, C. E. Motor neuron diversity in development and disease. *Annu Rev Neurosci* **33**, 409-440, doi:10.1146/annurev.neuro.051508.135722 (2010).

18 Jaitin, D. A. *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776-779 (2014).

19 Chung, C. Y. *et al.* Identification and Rescue of α-Synuclein Toxicity in Parkinson Patient–Derived Neurons. *Science* **342**, 983-987 (2013).

20 Trusina, A., Papa, F. R. & Tang, C. Rationalizing translation attenuation in the network architecture of the unfolded protein response. *Proc Natl Acad Sci U S A* **105**, 20280-20285, doi:10.1073/pnas.0803476105 (2008).

21 Heiman, M. *et al.* A translational profiling approach for the molecular characterization of CNS cell types. *Cell* **135**, 738-748 (2008).

22 Rothstein, J. D. & Diamond, B. The Role of NG2 Glial Cells in ALS Pathogenesis. (DTIC Document, 2013).

23 Di Giorgio, F., Boulting, G., Bobrowicz, S. & Eggan, K. Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. *Cell Stem Cell* **3**, 637-648, doi:10.1016/j.stem.2008.09.017 (2008).

24 Marchetto, M. C. N. *et al.* Non-Cell-Autonomous Effect of Human SOD1< sup> G37R</sup> Astrocytes on Motor Neurons Derived from Human Embryonic Stem Cells. *Cell Stem Cell* **3**, 649-657 (2008).

25 Lobsiger, C. S. & Cleveland, D. W. Glial cells as intrinsic components of non-cell-autonomous neurodegenerative disease. *Nat Neurosci* **10**, 1355-1360 (2007).

26 Gladman, M., Cudkowicz, M. & Zinman, L. Enhancing clinical trials in neurodegenerative disorders: lessons from amyotrophic lateral sclerosis. *Curr Opin Neurol* **25**, 735-742, doi:10.1097/WCO.0b013e32835a309d (2012).

27 Roybon, L. *et al.* Human stem cell-derived spinal cord astrocytes with defined mature or reactive phenotypes. *Cell Reports* **4**, 1035-1048 (2013).

28 Selvaraj, V., Jiang, P., Chechneva, O., Lo, U. & Deng, W. Differentiating human stem cells into neurons and glial cells for neural repair. *Frontiers in bioscience (Landmark edition)* **17**, 65-89 (2011).

29 Francius, C. & Clotman, F. Generating spinal motor neuron diversity: a long quest for neuronal identity. *Cellular and Molecular Life Sciences* **71**, 813-829 (2014).

30 Alaynick, W. A., Jessell, T. M. & Pfaff, S. L. SnapShot: spinal cord development. *Cell* **146**, 178 (2011).

31 Foo, L. *et al.* Development of a method for the purification and culture of rodent astrocytes. *Neuron* **71**, 799-811, doi:10.1016/j.neuron.2011.07.022 (2011).

32 Gould, T. *et al.* Complete dissociation of motor neuron death from motor dysfunction by Bax deletion in a mouse model of ALS. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **26**, 8774-8786, doi:10.1523/jneurosci.2315-06.2006 (2006).

33 Williams, A. H. *et al.* MicroRNA-206 delays ALS progression and promotes regeneration of neuromuscular synapses in mice. *Science* **326**, 1549-1554 (2009).

34 Wichterle, H., Lieberam, I. & Porter, J. Directed differentiation of embryonic stem cells into motor neurons. *Cell* (2002).

35 Boulting, G. L. *et al.* A functionally characterized test set of human induced pluripotent stem cells. *Nature biotechnology* **29**, 279-286 (2011).

36 Xu, C. *et al.* A Zebrafish Embryo Culture System Defines Factors that Promote Vertebrate Myogenesis across Species. *Cell* **155**, 909-921 (2013).

37 Benatar, M. Lost in translation: treatment trials in the SOD1 mouse and in human ALS. *Neurobiol Dis* **26**, 1-13, doi:10.1016/j.nbd.2006.12.015 (2007).